

Causality, Bell's theorem, and Ontic Definiteness

Joe Henson*

January 12, 2013

Abstract

Bell's theorem shows that the reasonable relativistic causal principle known as “local causality” is not compatible with the predictions of quantum mechanics. It is not possible maintain a satisfying causal principle of this type while dropping any of the better-known assumptions of Bell's theorem. However, another assumption of Bell's theorem is the use of classical logic. One part of this assumption is the principle of *ontic definiteness*, that is, that it must in principle be possible to assign definite truth values to all propositions treated in the theory. Once the logical setting is clarified somewhat, it can be seen that rejecting this principle does not in any way undermine the type of causal principle used by Bell. Without ontic definiteness, the deterministic causal condition known as Einstein Locality succeeds in banning superluminal influence (including signalling) whilst allowing correlations that violate Bell's inequalities. Objections to altering logic, and the consequences for operational and realistic viewpoints, are also addressed.

1 Introduction

Bell's theorem [1, 2, 3] encapsulates by far the most mysterious aspect of quantum mechanics, often summed up as “quantum non-locality”. Indeed, Spekkens has put forward an ingenious, but distinctly unmysterious, hidden-variable model that qualitatively reproduces practically every interesting feature of quantum mechanics (interference, no-cloning, *etc.*) *apart* from nonlocality [4]. The theorem demonstrates that a reasonable causal principle which Bell named “local causality” is incompatible with the predictions of quantum mechanics, given some other seemingly innocuous assumptions, chiefly “freedom of settings” from correlations with events in the past that are relevant to the experiment, and of course the use of standard logic. As has been noted by several authors, the assumption of local causality can be broken down into two parts, either of which can be rejected [5, 6]. The first is the assumption that there should be a Principle of Common Cause (PCC) [7], that is, that correlations

*jhenson@perimeterinstitute.ca

between events require a causal explanation. In this view, if event A is correlated to B , either the events are directly causally related, or the correlation can be explained by a “common cause” C (additionally to this, a cause must be to the past of its effect). The second is the principle of Relativistic Causal Structure (RCS), which states that the relativistic spatiotemporal causal structure should be used to define the past of any spacetime region, a position forced on us if we accept relativity¹.

This creates a difficult choice for anyone attempting to create a satisfying interpretation for quantum mechanics. At least one of Bell’s assumptions must be dropped. Approaches such as deBroglie-Bohm pilot wave theory [8, 9] and those based on Nelson’s mechanics [10, 11] drop the RCS assumption, whereas the GRW “spontaneous state-vector collapse” program [12] in Bell’s “flash” ontology [13], Kent’s “Real Worlds” program [14] and arguably the consistent histories program [15] (to the extent that it aims at producing a unique probability distribution of over a set of events) all drop the PCC assumption. Suggestions have been made that the freedom of settings assumption could be dropped instead (either in the guise of past events influencing setting choices, or of “backwards causation”) [16, 17].

These considerations lead us to the first main argument of this paper. Bell’s theorem creates the great conceptual difficulties it does for the following reason: it seems to preclude the possibility of recovering a satisfying causal principle in any theory that contains quantum mechanics. Each of the options currently on the table leaves something to be desired in this regard, as discussed at greater length below. This issue is distinct from the measurement problem. Leaving the measurement problem unsolved simply begs the question, as we will see, and is, at bottom, a particular way of dropping the PCC as an assumption. Below, it is maintained that a satisfying causal principle, given what we know of physics, should be relativistic. This at least means that it should ban superluminal signalling. As we will see, without any one of the three assumptions mentioned above, this cannot be achieved, and so the lack of superluminal signalling remains unexplained by a causal principle in interpretations which drop one of these – which is distinctly unsatisfying. However, a satisfying causal principle needs to be more than simply a ban on superluminal signalling, for reasons that are also addressed. Local causality, or its deterministic cousin, Einstein locality, are better examples of satisfying causal conditions.

Why should we want to reinstate such a causal principle at all? The reasons for this are treated at greater length in an appendix to this article. The most basic reason is that doing so may enhance our understanding of present theory, and increase theoretical progress. Also, if a relativistic version of the PCC could be made compatible with quantum mechanics, it would be a restrictive condition on theories. This could serve as a guiding principle in the search for new theories. For example, this idea has already been applied in the context of the causal set approach to quantum gravity, at least in a classical setting. In this case, a stochastic toy model was developed by taking a general dynamical framework and then selecting a physically appealing set

¹Thus, it is not accurate to say that Bell’s theorem shows a contradiction between a realistic world-view and relativity (given the other assumptions); the two can be made compatible by dropping the PCC, just as they are in operational interpretations [6].

of dynamics by applying basic physical principles, including a causal principle [18].

The one path left to us is the alteration of logic. Curiously, although the use of nonstandard logics has often been considered in the study of the foundations of quantum mechanics, this has never been motivated by the possibility of evading Bell’s theorem. Before the problem of Bell’s theorem became apparent, Birkhoff and von Neumann attempted to consistently incorporate all projection operators as propositions in one “quantum logic” [19, 20], while Reichenbach’s three-valued logic approach [21] sought first and foremost to give meaning to complimentary statements, and to complementarity itself, inside the logic being employed, rather than declaring some statements meaningless (although he did come close to some aspects of the discussion below in his treatment of EPR “causal anomalies”). Most recently, topos theory has been used to incorporate contextuality in a pseudo-realist setting [22, 23], whilst anhomomorphic logic was introduced to deal with interference in an observer-independent sum-over-histories account of quantum theory [24, 25, 26].

Altering logic to evade Bell’s theorem seems at first like a very drastic and unnatural step to take, which could only be motivated by desperation to avoid the difficulties of other approaches. Several questions immediately arise. Firstly, our causal principle of local causality, and freedom of settings, are built on the foundation of normal logic. At first glance, then, it seems that we would have to reject local causality with any attempt to alter logic – a case of throwing out the baby with the bathwater. Proposing a different causal principle in a different logical setting may, therefore, seem therefore like nothing more than an arbitrary proposal that could have been designed to reach the desired conclusion, rather like moving the goalposts in order to score. Also, several related worries may arise about the connection to the everyday world of standard logic. Can causation itself be given a satisfying meaning in any particular logical framework? And, when reasoning *about* the new logical system, are we not applying classical logic in any case? The justifications for, and ramifications of, taking so radical a step should be considered carefully.

We will see that these concerns can be answered in at least one case of altered logic. First, a third “intermediate” truth value is added between true and false. This can be done without changing anything in the standard picture, if we treat the indefinite truth value as signifying ignorance of the “real”, definite, truth value. The break with standard logic comes when we question whether there need be a definite truth value for all propositions in all circumstances. A principle along these lines is dubbed *ontic definiteness* below. As we will see, rejecting this assumption does *no damage at all* to the definition or meaning of Einstein locality or the freedom of settings; together they still ban superluminal signalling, for instance. Even so, this move does allow us to evade Bell’s theorem. Dropping ontic definiteness thus allows us to preserve a satisfying causal principle without running into a contradiction with the predictions of quantum mechanics. Rather than being a desperate measure forced on us by Bell’s theorem, it is arguable that this innovation would have been welcomed by many philosophers and mathematicians independently of the problem of Bell’s theorem. Besides this – whether or not this should affect our judgements – it is not as far from common sense as many “quantum logics” that have been proposed.

The structure of the paper is as follows. In section 2, a framework is given in which a relativistic causal principle can be defined, as an example of what is being searched for. Section 3 discusses the consequences of dropping assumptions of Bell’s theorem for this type of principle. In section 4, the assumption of standard logic is addressed, and the assumption of ontic definiteness is defined, leading to a discussion of how quantum correlations in the EPRB experiment may be given a causal explanation. A discussion of the consequences of this view, and further possible objections, is given in section 5. The appendices deal with justifications for insisting on a satisfying causal principle (appendix A), relations between various causal principles (appendix B), and a comparison to Reichenbach’s three-valued logic (appendix C).

2 A Causal Principle

Before explaining anything further about maintaining a satisfying causal principle when faced with Bell’s theorem, it is worthwhile to consider why we would want to do so. Since it is somewhat outside the main topic of the article, comments on this topic are relegated to Appendix A. All that is necessary here is to make the following modest assertion: finding a satisfying causal principle that is compatible with quantum theory may help us to make scientific progress. The intuition behind this will be clarified in the process of defining a causal principle and examining the issue in the context of Bell’s theorem.

The next question is: what exactly is meant by a satisfying causal principle? We can now turn to a discussion of a good example².

2.1 Einstein locality

Let us first consider a deterministic, relativistic principle of causality, Einstein Locality (a term apparently introduced by Bell and d’Espagnat [30]). Discussion will be limited, for the remainder of this article, to theories with fixed background spacetime. EL is a restriction on non-probabilistic theories, which we can define in the following way. First, we consider a space of histories Ω . Each history is supposed to represent a full specification of all physical properties of the system being modelled at all times. Events are subsets of Ω . The word “event” here is taken to mean something that may or may not occur according to the theory, *i.e.* events correspond to propositions in the theory, and the subset of Ω is the set of histories for which the proposition is true. There is a Boolean algebra of events Σ , which contains all the events that we can ask about in this theory³. This simply means that for questions about pairs of events,

²It is worth noting here that, despite some discussion, expressions of the type put forward by Reichenbach and Bell have been convincingly defended from various claims that principles of common cause might reasonably be made significantly weaker [5, 27, 28, 29]. See also the comments at the end of the following section.

³For the purposes of this article it will be convenient to ignore the complications that arise for history spaces with uncountable cardinality, in which case the algebra should more properly be a Boolean sigma-algebra.

logical operations are defined in the normal way. Each history $\gamma \in \Omega$ corresponds to a truth valuation on this algebra, technically a homomorphism $\gamma : \Sigma \rightarrow \{0, 1\}$, where $\{0, 1\}$ is the two element Boolean algebra. Here, 0 is taken to mean “false” and 1 “true”. For instance, event A might signify that there is a ball in a particular box, and if $\gamma(A) = 1$, that would be true for the history γ . That $\gamma(\cdot)$ is a homomorphism means that the usual rules of logic apply to the truth values, for instance, if “ A ” is false then “not A ” must be true. We will use set theoretic notation so that $A \cap B$ means, in words, “ A and B ”, A^c (the set complement) means “not A ”, and so on.

We also need to introduce the background spacetime, which we call \mathcal{M} . For our purposes this \mathcal{M} can be defined as a weakly causal Lorentzian manifold⁴ (although the formalism could be just as easily applied to a discrete causal set). To say that our spacetime is equipped with a spatiotemporal causal structure means that there is a partial order \prec on its points, where $x \prec y$ signifies that point x is to the past of point y . We call subsets of the spacetime “regions”. For a region $\mathcal{R} \subset \mathcal{M}$, the “causal past” $J^-(\mathcal{R})$ is defined to be all those points that are to the past of (or equal to) any point in \mathcal{R} .

Each event occurs somewhere in our fixed background spacetime. Thus, we define a map between regions and subalgebras of Σ , $\Delta : \text{pow}(\mathcal{M}) \rightarrow \text{pow}(\Sigma)$, where $\text{pow}(\mathcal{M})$ is the set of all subsets of \mathcal{M} and $\text{pow}(\Sigma)$ is the set of all subalgebras of Σ . That is, to each region \mathcal{R} we assign a Boolean subalgebra of Σ , which we label $\Sigma|_{\mathcal{R}}$, and we say that an event $A \in \Sigma|_{\mathcal{R}}$ is *associated to* the region \mathcal{R} . This amounts to saying that, if a set of events is associated to a region, then logical combinations of those events should also be associated to that region. These “regions of association” must obey certain consistency conditions which we need not go into in detail here (the reader is referred to [29]). It will suffice to say that if $\mathcal{R} \subset \mathcal{R}'$ then $\Sigma|_{\mathcal{R}}$ is a subalgebra of $\Sigma|_{\mathcal{R}'}$. Truth valuations can be restricted to any subalgebra in the obvious way, and we will denote the restriction of $\gamma(\cdot)$ to the subalgebra associated to \mathcal{R} as $\gamma|_{\mathcal{R}} : \Sigma|_{\mathcal{R}} \rightarrow \{0, 1\}$. The collection of all such restrictions $\gamma|_{\mathcal{R}}$ of truth valuations corresponding to histories $\gamma \in \Omega$ will be called $\Omega|_{\mathcal{R}}$.

A non-probabilistic theory on the history space also provides a subset $\Theta \subset \Omega$ that contains the histories allowed by the theory. For instance, if Ω was the set of scalar field configurations $\phi(x)$, Θ might contain only the solutions to a particular field equation. We can imagine the theory as a pack of cards, each representing a history in Θ . If, by experiment and observation, we know the truth values of a certain set of events, we can go through the deck, keeping the cards that agree with our knowledge and throwing the rest away (events whose truth values we don’t know are allowed to be true or false on the cards that we keep). If we have no cards left, our theory has been falsified. Otherwise, we can use the remaining cards to make inferences about other events (including predictions), if they have the same truth value on all of those cards.

We will call two events “correlated” (resp. “anticorrelated”) if and only if they share the same truth value (resp. have opposite truth values) for all histories in Θ .

⁴This is a differentiable manifold equipped with a Lorentzian metric that admits no closed causal curves, allowing a well-defined causal relation between all pairs of points.

Another way to state this is that A and B are correlated if and only if $A \cap \Theta = B \cap \Theta$. We will say that A is *weakly correlated* to B if and only if $A \cap \Theta \subset B \cap \Theta$, so that $\gamma(A) = 1$ implies that $\gamma(B) = 1$ for all histories in Θ . Weak correlation includes the case of full correlation. Finally, if we want to discuss probabilities for events, we can put a probability distribution on Ω that only has support in Θ .

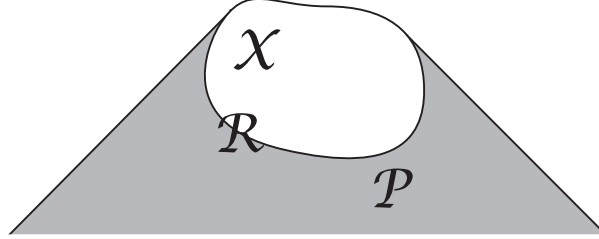


Figure 1: A “spacetime diagram” of the regions involved in the definition of Einstein Locality (here \mathcal{M} is represented as 2D Minkowski space in a usual spacetime diagram, with time running up the page). The region \mathcal{R} is the union of the unshaded region \mathcal{X} and the shaded past region $\mathcal{P} = J^-(\mathcal{X}) \setminus \mathcal{X}$ within (and including) the light-cone. According to EL, the truth values of events in \mathcal{P} must determine the truth values of all events in \mathcal{R} .

We now have enough machinery in place to define Einstein Locality. Consider a region \mathcal{X} , its past $\mathcal{R} = J^-(\mathcal{X})$ and its “exclusive past” $\mathcal{P} = J^-(\mathcal{X}) \setminus \mathcal{X}$ where \setminus is set difference, as shown in figure 1. The idea is that what happens in \mathcal{R} is determined by what happens in \mathcal{P} ⁵. So, if we fix the truth value of every event that occurs in \mathcal{P} , then this should fix all truth values in \mathcal{R} .

Einstein locality: Consider a nonprobabilistic theory defined by $\{\Omega, \Sigma, \Theta, \mathcal{M}, \Delta\}$, and regions \mathcal{X} , $\mathcal{R} = J^-(\mathcal{X})$ and $\mathcal{P} = J^-(\mathcal{X}) \setminus \mathcal{X}$. The theory is *Einstein Local* if the following condition holds for all regions \mathcal{X} :

$$\forall \gamma_{\mathcal{P}}^* \in \Omega|_{\mathcal{P}}, \exists \gamma_{\mathcal{R}}^* \in \Omega|_{\mathcal{R}} \text{ such that } \forall \gamma \in \Theta, \text{ if } \gamma|_{\mathcal{P}} = \gamma_{\mathcal{P}}^* \text{ then } \gamma|_{\mathcal{R}} = \gamma_{\mathcal{R}}^*.$$

That is, if only we knew everything about events in \mathcal{P} , we could infer the truth value of all events in \mathcal{R} . That is what is meant by events in \mathcal{P} determining events in \mathcal{R} . To explain this principle let us return to the analogy of the deck of cards. We wish to keep in our hand only those cards that are possible histories. Einstein Locality says that, if we condition on all events associated with the past region \mathcal{P} , so that we only hold cards with a particular set of truths values for these (called $\gamma_{\mathcal{P}}^*$), then we will find that, for all events in \mathcal{R} , the same truth values (given by $\gamma_{\mathcal{R}}^*$) will be given on every remaining card.

There is another way of stating Einstein Locality:

⁵The reason that we use the region \mathcal{R} rather than \mathcal{X} is to make our condition sound even if we allow that some events associated with \mathcal{R} may not be logical combinations of events associated with \mathcal{X} and events associated with \mathcal{P} . This type of condition is known as “separability of events”, and is called property (iv) in [29].

Einstein Locality (second formulation): Consider a nonprobabilistic theory defined by $\{\Omega, \Sigma, \Theta, \mathcal{M}, \Delta\}$, and regions \mathcal{X} , $\mathcal{R} = J^-(\mathcal{X})$ and $\mathcal{P} = J^-(\mathcal{X}) \setminus \mathcal{X}$. The theory is *Einstein Local* if the following condition holds for all regions \mathcal{X} :

$$\forall A \in \Sigma|_{\mathcal{R}}, \exists B \in \Sigma|_{\mathcal{P}} \text{ such that } \gamma(A) = \gamma(B) \forall \gamma \in \Theta.$$

In words, this means that for all events associated to \mathcal{R} there is an event associated to \mathcal{P} that is correlated to (or “determines”) it. The determining past event B will sometimes be termed the *causal antecedent* of A below. The proof that these two definitions are equivalent is relegated to appendix B.1.

This condition has a close link to Reichenbachian principles of common cause [7, 29]. Einstein locality encapsulates a nonprobabilistic PCC (NPCC). The following is a “joint past” version of this principle:

NPCCj: Consider a nonprobabilistic theory defined by $\{\Omega, \Sigma, \Theta, \mathcal{M}, \Delta\}$, two space-like regions \mathcal{A} and \mathcal{B} , and their “joint past” $\mathcal{P}_j = J^-(\mathcal{A}) \cup J^-(\mathcal{B}) \setminus (\mathcal{A} \cup \mathcal{B})$ as shown in figure 2. The theory obeys the NPCCj if the following principle holds: for any pair of events $A \in \Sigma|_{\mathcal{A}}$ and $B \in \Sigma|_{\mathcal{B}}$, if A and B are correlated, then there exists an event $C \in \Sigma|_{\mathcal{P}_j}$ that is correlated to both B and C .

This follows from the second formulation of Einstein locality: that principle requires that event there is an event C in the past (that is, associated to $J^-(\mathcal{A})$, which is a subset of the joint past \mathcal{P}_j) such that $\gamma(A) = \gamma(C) \forall \gamma \in \Theta$, and if $\gamma(A) = \gamma(B) \forall \gamma \in \Theta$, then C is also correlated to B . By considering the complement of events we can easily extend this reasoning to anticorrelated events as well. There is also a “mutual past” version of this principle:

NPCCm: Consider a nonprobabilistic theory defined by $\{\Omega, \Sigma, \Theta, \mathcal{M}, \Delta\}$, two space-like regions \mathcal{A} and \mathcal{B} , and their “mutual past” $\mathcal{P}_m = J^-(\mathcal{A}) \cap J^-(\mathcal{B})$ as shown in figure 2. The theory obeys NPCCm if the following principle holds: for any pair of events $A \in \Sigma|_{\mathcal{A}}$ and $B \in \Sigma|_{\mathcal{B}}$, if A and B are correlated, then there exists an event $C \in \Sigma|_{\mathcal{P}_m}$ that is correlated to both B and C .

These two versions on the NPCC are in fact equivalent, as is proven in appendix B.2. Thus, both of them follow from Einstein Locality.

What is the relation between Einstein Locality and other causal principles? In the case of a non-relativistic theory, instead of a relativistic spatiotemporal causal structure, we simply put all events that are associated to spacetime regions before time t to the past of all events that are associated to regions after or at that time. Above, substituting the relativistic causal structure of our Lorentzian manifold \mathcal{M} with this kind of causal structure would give the kind of determinism found in Newtonian theories (sometimes called Millsian causation [31]).

One might wonder if it is necessary to have strict determinism to have a good causal principle that incorporates the PCC and bans superluminal influence. Indeed, for this purpose is not necessary to have everything in \mathcal{R} determined by events in

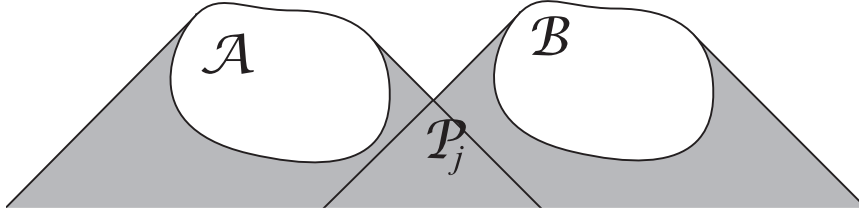


Figure 2: A “spacetime diagram” of the regions involved in the definition of the Nonprobabilistic Principle of Common Cause (joint past version). The spacelike regions \mathcal{A} and \mathcal{B} are shown unshaded, while $\mathcal{P}_j = J^-(\mathcal{A}) \cup J^-(\mathcal{B}) \setminus (\mathcal{A} \cup \mathcal{B})$ is the entire shaded region between (and including) both lightcones. According this PCC, any correlated pairs of events in regions \mathcal{A} and \mathcal{B} must be correlated to a further event in \mathcal{P}_j .

\mathcal{P} , but only necessary that, once we have conditioned on all events in \mathcal{P} , nothing more about the truth values of event in \mathcal{R} can be determined by conditioning on any events that occur spacelike to \mathcal{R} . This line of argument leads to a weaker condition. However (and although giving a thorough defence of this would require a separate article), I claim that (a) this weakening does not evade Bell’s theorem and (b) we will be able to maintain the stronger condition of Einstein locality when we alter logic. Because of this, the weaker condition is not relevant for the following discussion.

Beyond this, there are the probabilistic formulations such as Bell’s local causality. It is not hard to convince oneself that these follow from Einstein locality in the following sense. If we take the histories in Θ and put a probability distribution over them, then, if the original theory was Einstein local, the resulting probabilities theory will obey a version of local causality called SO2 in [29]. This derivation is left for appendix B.3. Again, it is well known that using the probabilistic version does not evade Bell’s theorem. There are several other formulations of probabilistic, relativistic causal principles, but in previous work it was conjectured that all of these are either problematic, equivalent to local causality, or incorporate additional inessential assumptions [29], and steps have been taken towards showing this (see also [27]). The conclusion, again, is that Einstein locality is a reasonable condition to concentrate on for the purposes outlined in the introduction of this article. The important thing to keep in mind is that Einstein locality expresses the principle that there are no superluminal influences, as all events are determined by events in their past light-cone. It encompasses a combination of a principle of common cause, and relativistic causal structure. In particular, when the theory describes agents and devices, Einstein locality bans superluminal signalling.

2.2 Einstein locality and superluminal signalling

As explained in appendix A.2, a distinction needs to be drawn between a simple operational ban on superluminal *signalling*, and causal conditions like EL which ban superluminal *influence*. Nonetheless, since signalling is a type of influence, any rela-

tivistic causal condition worth its salt should imply a ban on superluminal signalling. Fortunately it is not hard to see that EL does this.

We can now give a straightforward account of how Einstein locality gives rise to a ban on superluminal signalling. Consider an agent in a region \mathcal{A} trying to influence something in a spacelike region \mathcal{B} . Some event $A \in \Sigma|_{\mathcal{A}}$ is to be regarded as the input event, and some event $B \in \Sigma|_{\mathcal{B}}$ will be the output. In the context of non-probabilistic theories set out above, we can distinguish *strong* and *weak* signalling. Let us consider strong signals first. A strong signal is successful if and only if A and B are correlated, which as before means that $\gamma(A) = \gamma(B) \forall \gamma \in \Theta$. This is not sufficient to define signalling, however. For example, if two newspapers delivered to two different houses have the same headline, this does not constitute a signal between the newspapers. Both events were determined by a common cause in the past. Thus we must impose some condition to make sure that event A is a “free setting”: that is, A is free from the influence of past events that may also influence B . One requirement is that, in our model, A is not determined by any event in its past region $J^-(\mathcal{A}) \setminus \mathcal{A}$. If we are applying EL, we must in this case exempt the event A from being determined, basically because we have left out its causal antecedents from our model. For instance, we might decide A based on what we had for breakfast. “Freedom of settings” means assuming that what we had for breakfast, or any other such factor, is irrelevant when modelling the system in question. The event B however is not free. With A “free” in this sense, if $\gamma(A) = \gamma(B) \forall \gamma \in \Theta$, then we must admit that there is superluminal signalling from \mathcal{A} to \mathcal{B} .

It can now be seen that this can be ruled out by applying EL. Consider the region $\mathcal{Z} = J^-(\mathcal{B}) \setminus J^-(\mathcal{A})$. Because $\mathcal{B} \subset \mathcal{Z}$, it follows that $B \in \Sigma|_{\mathcal{Z}}$, and so by Einstein Locality B requires a causal antecedent in $J^-(\mathcal{Z}) = J^-(\mathcal{A}) \cap J^-(\mathcal{B})$, that is, an event C such that $\gamma(B) = \gamma(C) \forall \gamma \in \Theta$. But if $\gamma(A) = \gamma(B) \forall \gamma \in \Theta$ this implies $\gamma(A) = \gamma(C) \forall \gamma \in \Theta$ which contradicts our freedom of settings assumption for A , because C is associated to $J^-(\mathcal{A}) \cap J^-(\mathcal{B})$ and thus is associated to the past of \mathcal{A} ⁶.

Weak signalling means that the truth value of the signal *can* affect the value of the output, but does not in all circumstances. There is a weak signal between A and B if and only if A is a free setting, and B is weakly correlated to A . In this case, $\gamma(B) = 0$ whenever $\gamma(A) = 0$ for all histories in Θ , and so the truth value of A can determine the truth value of B , although the events may not be fully correlated in the sense already defined. Changing the setting could thus affect the truth value of the event B . The definition could just as easily have been that A is weakly correlated to B (rather than B to A), but this is a matter of convention if A^c is also regarded as a signal, which it will be.

This line of thought also shows that our condition for an event like A to be a free setting is incomplete: we do not want any event C in $J^-(\mathcal{A}) \setminus \mathcal{A}$ to be even weakly

⁶This conclusion might seem surprising: why cannot the chain of causal antecedents of B remain outside of the past of \mathcal{A} arbitrarily far into the past, or until some initial cosmological hypersurface? Such a scenario would imply that the setting event was correlated to some initial cosmological event outside its past light-cone, or events arbitrarily far in the past, however, and it is reasonable that freedom of settings and EL together ban this. A similar argument is dealt with in appendix B.2 (see [29] for more on this point).

correlated to A . If it were, $\gamma(A) = 1$ is forced on us if we have $\gamma(C) = 1$, and so A can hardly be described as free from influence from its past. Again, Einstein Locality ensures that there can be no superluminal signalling, even of this weak variety. We previously established that, if EL holds, there is an event C associated to the past of \mathcal{A} that is correlated to B . If B is weakly correlated to A , then C must also be weakly correlated to A , violating our strengthened freedom of settings condition.

In the probabilistic setting, freedom of settings again means that the setting A should not be probabilistically correlated to any event in the past region (and neither should it become so after we condition on any other event in the past). Here we are using the word correlation in the usual probabilistic sense. This is the assumption that goes into Bell’s theorem.

3 Dropping assumptions of Bell’s theorem

Having discussed an example of a satisfying causal principle, we now examine the effects on the maintenance of such a principle if we drop assumptions of Bell’s theorem.

3.1 Dropping counterfactual definiteness?

It has been claimed by some that, as well as local causality, Bell’s theorem depends on a certain “hidden-variables” assumption. Thus those who object to the lack of local causality in Quantum Mechanics suggested by Bell’s theorem, it is argued, need only reject the idea of hidden variables. This view is summed up by Mermin [32] when he says “to those for whom nonlocality is anathema, Bells Theorem finally spells the death of the hidden-variables program.”

For simplicity let us consider measurements with outcome valued in $\{0, 1\}$ (*i.e.* measurements corresponding to single events). Counterfactual definiteness asserts that, for every possible measurement, there must be an event (occurring in the same region as the measurement) that shares the truth value that *would have* been obtained if that measurement was performed, irrespective of whether the measurement actually was performed. This is sometimes cited as an assumption of Bell’s theorem, leading to conclusions like Mermin’s. One interpretation of the sentence quoted above is that there is a choice between locality and counterfactual definiteness, and that, in operational interpretations in particular, we can have a satisfactory resolution to the problem of Bell’s theorem if we drop counterfactual definiteness.

This interpretation is mistaken, however. Norsen (quoting other earlier treatments) convincingly argues that counterfactual definiteness is *not* a separate assumption, but in fact follows from local causality (and the results of quantum mechanics which specify that perfect correlations between some outcome events can be achieved in the EPRB set-up) [33]. Indeed, it is clear in Bell’s last published account of his result [3] that the argument begins with causality and leads to counterfactual definiteness. Thus, whether or not we start with a bias on the subject of hidden variables, if we accept Bell’s local causality as a good formalisation of the idea of a relativistic causal principle, then we are forced to accept counterfactual definiteness. In other

words, when standard logic is assumed, rejecting counterfactual definiteness amounts to rejecting local causality (and Einstein Locality).

The lack of superluminal signalling is often mentioned in the operational context, to justify quantum theory against the principle of relativity. Be that as it may, lack of superluminal signalling cannot be substituted for local causality; it is not in itself a good relativistic causal principle of the type being searched for here, as argued in appendix A.2. The problem of quantum nonlocality still remains. Thus, “to those for whom nonlocality is anathema”, *all* of the options presently on the table should be equally unsatisfying.

The conclusion is that there is no separate assumption of Bell’s theorem here. Despite this, later we will see that there is a way to resurrect the counterfactual definiteness “loophole” while maintaining Einstein Locality. This is made possible by dropping another, deeper assumption.

3.2 Dropping the PCC

Dropping the principle of common cause itself clearly amounts to throwing in the towel as far as maintaining a causal principle is concerned. In this approach, there are still prospects for doing away with measurement as a fundamental notion in quantum mechanics, but one gives up on another goal of quantum interpretation: that of inferring facts about the quantum world using a causal principle, as briefly argued in appendix A. In interpretations like Kent’s Real Worlds [14], and consistent histories [15](in one reading), the goal is to extract from the quantum formalism some probability distribution on a coarse-grained algebra of events, which should at least be larger than any set of questions that we could ever answer operationally. But confronted with the strange correlations of the EPRB experiment, one simply claims that correlations do not need to be explained. Although the production of the original Bell state in the past of both wings of the experiment seems to be the “cause” in some sense, this has not been formalised. But how, in this view, do we explain the lack of superluminal signalling in relativistic quantum theories? It can have nothing to do with a deeper lack of superluminal influence, talk of which we have forsworn; it is merely an unexplained fact. GRW “state-vector-collapse” models [12], most clearly in Bell’s flash ontology [13], have the same feature.

Operational treatments of quantum mechanics should also be put in this category. In this case the events that we allow in our theory are only those in some set of “operational” events, referring explicitly to preparations and observations. But nothing stops us from conceiving of such theories as probability distributions over this rather impoverished algebra of events (albeit that the set of events is not given by the theory, but defined by fiat, due to the arbitrary split between the classical and quantum worlds), which can as above be located in spacetime regions. Again, we find that the PCC is not respected; correlations between spacelike events cannot always be explained by conditioning on past events. As Norsen has noted, then, there is something wrong with the assertion that the assumptions of Bell’s theorem can be summed up as “local realism”, as if there is a separate assumption of realism involved, the rejection of which relieves us of the conundrum of Bell’s theorem [34]. Rather,

operationalist accounts also have to drop the assumption of the PCC, and thus give just as bad an account of causality as some of the so-called “realist” alternatives.

I will also claim that many-worlds interpretations should be placed in this category and thus fail to maintain a satisfying causal principle. The point has been made by Timpson and Brown [6] that the PCC has to be dropped in this case for consistency with relativity⁷.

3.3 Dropping RCS

Instead we can insist on the PCC, and freedom of settings, but reject relativistic causal structure. In doing so we can at least maintain a non-relativistic, “Millsian” version of causation as in Newtonian theories. This is the position taken in deBroglie-Bohm pilot wave theories [8, 9], or in reworkings [11] of Nelson’s “stochastic mechanics” programme [10], for example. In this view, the lack of superluminal signalling is merely a peculiarity of the theory, which is not related to any objective lack of superluminal influence. Likewise the apparent relativistic nature of events at the classical level is just an emergent peculiarity that is not reflected in the fundamental theory.

The following discussion will be based on the position that this does not add up to a satisfying causal principle. We have added extra events to quantum mechanics in order to satisfy the PCC. But we do not find an ontology that agrees with important physical principles like those of relativity. Neither do we recover a causal principle that explains the lack of superluminal signalling. Instead, we end up with experimentally inaccessible events that violate relativity and allow superluminal influences. It is a reasonable (though not inescapable) conclusion that something has gone wrong: “to those for whom nonlocality is anathema”, these hypothetical new events provide an unlikely explanation of apparent events, and so are not telling us anything useful. The new underlying variables seem to be a “fifth wheel” in the theory.

3.4 Dropping Freedom of Settings

This leaves the possibility of dropping the “freedom of settings” assumption. This seems to provide us with a way to keep both a causal principle and relativity in the face of Bell’s theorem. Given the “apparent” events (settings and outcomes) described by Bell, we can “fine-grain” the description, adding hypothetical events to the past in order to find a model that contains the apparent events, but which also obeys local causality. The problem with this approach is that *any* probability distribution over events happening in spacetime can be “causally completed” in this way. This follows from a theorem of Placek’s [35]⁸. This includes even the most

⁷However, the following discussion in [6] attempts to give a causal account of the EPRB experiment in the many-worlds setting in a different way. These ideas seem to me inadequate to recover a satisfying causal principle in the sense that it is being searched for here, however. Discussion of these subtleties will be left for future work.

⁸The theorem does not consider the issue of *where* in spacetime events occur in the same way as has been done in the present work, but this issue can easily be dealt with by putting the added

blatant cases of superluminal signalling. For example, as in the discussion of section 2.2, our agent could be controlling a “setting” event A in a region \mathcal{A} spacelike to region \mathcal{B} , containing event B . Let us say that the events are perfectly correlated. Without the freedom of settings requirement, we can make a causal model of this simply by adding an event C in the past that is perfectly correlated with A and B . Perhaps C could describe some event involving the decay of a “conspiritor” particle, one half of which travels off to determine B whilst the other half goes off to determine the agent’s decision on what value of A to choose – and also anything correlated to her decision, for instance what she had for breakfast, if she chooses the setting on that basis.

Without any further restriction, the proposed relativistic causal principle is completely empty. In particular, it allows superluminal signalling. In this case it is not deserving of the name; it is not a satisfying relativistic causal principle. It is possible that, if we first imposed other conditions on our theory, local causality could regain some power, but then the problem would be justifying calling this set of conditions a causal principle. As well as banning superluminal signalling, a causal principle should be justifiable on counterfactual grounds. Although none of these considerations rule out making an interpretation of quantum mechanics based on this loophole, they do critically undermine motivation for doing so rather than pursuing other methods⁹.

As we have seen, whether or not to solve the measurement problem is only one of the major decisions that must be made when interpreting quantum mechanics. A second, perhaps equally important choice, is which assumption of Bell’s theorem to sacrifice. The most well-known interpretations all choose to drop either the PCC or relativistic causal structure. Explicitly or implicitly, this choice depends on whether or not we feel that the principle of common cause is so important to physics that it must be preserved, even at the expense of relativity. We have also seen that dropping any of the three assumptions that are most usually discussed leads to an unsatisfying situation regarding relativistic causal principles. The question now is, can we do better by exploiting the forth assumption, that of standard logic?

4 A modest alteration in logic

4.1 Possible objections

There are (at least) four objections that may cause some reticence to alter logic when faced with the problem of Bell’s theorem. Objection (1), “throwing the baby out

“hypothetical” events into the past of all of the original “apparent” events.

⁹Two other “loopholes” should be mentioned at this point. The first is that commented on by Morgan: that Bell correlations could be explained if we allow “initial correlations”, i.e. correlations between event on some initial spacelike surface before the experiment, or correlations that somehow propagate in from past infinity [17]. This is technically distinct from freedom of settings, but the arguments against exploiting it are very similar. Secondly there is “backwards causation”, as exploited in Cramer’s transactional interpretation [16], which can be seen as a different gloss on the rejection of freedom of settings from correlations with past events, and fails to maintain a satisfying causal principle for the same reason.

with the bathwater”, is that all the other assumptions of Bell’s theorem are premised on standard logic, and so it would seem that rejecting logic does not allow us to save them. In answer to this, we will see below that one particular part of normal logic can be removed without doing any damage to a particular formulation of Einstein locality. Objection (2), “moving the goalposts”, is similar, but directed at the *motivation* for the causal principle after logic has been altered. The argument is that the whole idea of causality depends on standard logic, and anything that might be called a causal principle in a non-standard logical setting is in fact an arbitrary invention. To combat this, in the details below it will be shown that such a principle can indeed be satisfying in the sense of banning superluminal influences (including signalling) and having a counterfactual motivation that is practically identical to that of Einstein Locality in the case of standard logic, giving it every bit as much meaning. Objection (3), “common sense”, is that logic should not be altered on empirical grounds (for debate on this point see [36, 37]); since it applies to the way we reason at all times – even while considering altering logic – and to everyday events, standard logic simply cannot be rejected. Hopefully, the modest alteration proposed below will not draw so much criticism in this regard, for the following reasons. Firstly, similar things have been proposed outside of the present context, and secondly, the usual reasoning for everyday or scientific purposes is not threatened. Objection (4), “the meta-level” objection, is that even if we had a physical theory based on non-standard logic, we would still have to reason *about* it using standard logic. Would it not always be possible to reduce the theory to one based on standard logic, and in that case, are we not stuck with the choice between the original three assumptions of Bell’s theorem? The details of this objection, and the answer to it, are somewhat more involved, and so discussion of this point will be deferred until we have an example to work with.

Standard logic (which we are applying to events in spacetime) consists of the Boolean algebra of events Σ , the standard space of truth valuations $\{0, 1\}$ and the standard type of truth valuations or “histories” γ , which are homomorphisms. Below, the Boolean algebra of events will be kept, but the space of truth valuations will be replaced with $\{0, 1/2, 1\}$. Reichenbach also considered such a move as a response to quantum mechanics [21]; the approach here differs significantly from his, as explained in Appendix C. The three truth values will first be introduced in a way that does no injury to standard ideas of logic¹⁰.

4.2 Introducing the intermediate truth value: “a theory of local knowables”

The Boolean algebra Σ of events has already been introduced, along with the idea of a truth valuation or history $\gamma : \Sigma \longrightarrow \{0, 1\}$, and the history space Ω . The histories γ represent possible realities here; we can talk as if one history is the one that actually exists, independent of our knowledge of the truth values of the various events. Histories like this can be termed *ontic* [4]. But what if we don’t know some

¹⁰The approach taken here could be called a simple type of “epistemic logic” [38], and also turns out to have similarities to Rescher and Brandom’s “schematic worlds” [39].

of the truth values? To put it another way, we might know that the ontic history lies in a certain subset of Ω , but have no knowledge about it beyond that. This gives rise to the idea of an *epistemic* history, which describes the state of knowledge rather than the most detailed possible description of the history. An easy way to describe an epistemic history is by a subset of Ω . But we can also describe it as a new kind of truth valuation on the algebra of events Σ . Here we need the extra truth value $1/2$, which we will call “intermediate”¹¹. “Three-valued truth valuations” will be truth valuations using the three values. For the purpose of this section (that is, to motivate the new logic), the term “three-valued” can be replaced with “epistemic”, and “intermediate” with “unknown”, without disturbing the meaning. These three-valued truth valuations can be derived from a set of ontic histories in the following way.

The rule for deriving a three-valued truth valuation $\tilde{\gamma}_X$ from a set of ontic histories X is as follows: if there exists some $t \in \{0, 1\}$ for which $\gamma(A) = t \forall \gamma \in X$, then $\tilde{\gamma}_X(A) = t$, otherwise $\tilde{\gamma}_X(A) = 1/2$. An equivalent way of saying this is that $\tilde{\gamma}_X(A) = 1$ if and only if $A \supset X$, $\tilde{\gamma}_X(A) = 0$ if and only if $A^c \supset X$, and $\tilde{\gamma}_X(A) = 1/2$ otherwise. This is exactly the intuitive rule: if we know the history is in some set for which our proposition is always true (resp. false), then the proposition must be true (resp. false); if, on the other hand, all we know is that the history is in some set for which our proposition can be true *or* false, then we don’t know its actual truth value. We will say that $\tilde{\gamma}_X$ is the three-valued truth valuation derived from the set X . Note that X is itself an event, and the events that are true for $\tilde{\gamma}_X$ are those that are logically implied by X . Note also that the first formulation of “deriving a three-valued truth valuation” given here can also be applied to a set of three-valued truth valuations, rather than ontic ones. This can be used to further coarse-grain the state of knowledge.

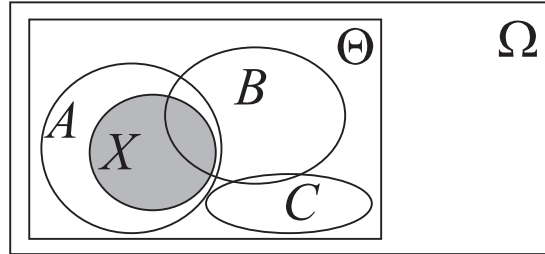


Figure 3: A Venn diagram depicting a history space Ω . Each point in the space represents an “ontic” history $\gamma \in \Omega$, and the subset Θ represents all histories allowed by our theory, as explained in section 2.1. A three-valued truth valuation $\tilde{\gamma}_X$ can be derived from the set X as explained in the text. It can be imagined that it is only known that the “real” history is in the set X . In the case shown, $\tilde{\gamma}_X(A) = 1$ because $\gamma(A) = 1$ for all $\gamma \in X$. In contrast, $\tilde{\gamma}_X(B) = 1/2$ and $\tilde{\gamma}_X(C) = 0$.

¹¹This term is meant to be considered philosophically neutral, rather than implying an “ontologically distinct” third truth value (which we will term “indefinite”) or a merely epistemic notion (which we will term “unknown”).

What is the logic that applies to this new kind of truth valuation? For definite truth values (meaning 0 and 1), standard truth tables apply. For intermediate values, there is not always a unique output for a logical operation. For instance, consider the operation “and”. If $\tilde{\gamma}(A) = \tilde{\gamma}(B) = 1/2$, then $\tilde{\gamma}(A \cap B)$ can equal 1/2 or 0. It cannot equal 1 as this would imply that both A and B must have the definite value 1. It is possible to set $\tilde{\gamma}(A \cap B) = 0$ however since this does not fix the ontic truth value of either A or B . This logic will not be examined in detail here; it is a sufficient definition that all three-valued truth valuations can be derived from a subset of some ontic history space Ω in the above sense.

What is a theory now? In the standard case, it was described by $\{\Omega, \Sigma, \Theta, \mathcal{M}, \Delta\}$. Each member of Ω is a history γ with an associated two-valued truth valuation $\gamma(\cdot)$. In the three-valued case this suggests $\{\tilde{\Omega}, \Sigma, \tilde{\Theta}, \mathcal{M}, \Delta\}$, where $\tilde{\Omega}$ is the space of all subsets of Ω , each of which has an associated three-valued truth valuation, and $\tilde{\Theta}$ is the subset of these that are also subsets of Θ . That is, if our three-valued theory is merely to be the epistemic version of the theory $\{\Omega, \Sigma, \Theta, \mathcal{M}, \Delta\}$, then $\tilde{\Theta}$ should just be the set of all three-valued truth valuations that can be derived from subsets of Θ . In this case we will say that $\tilde{\Theta}$ has been derived from Θ . When $\tilde{\Theta}$ is derived in this way, our ontic theory tells us that the “real” history is in the set Θ , and the histories allowed in our three-valued theory just express ignorance about which one it is. Epistemic theories $\{\tilde{\Omega}, \Sigma, \tilde{\Theta}, \mathcal{M}, \Delta\}$ derived in this way from ontic theories therefore obey the following principle¹²:

Ontic definiteness: A three-valued theory $\{\tilde{\Omega}, \Sigma, \tilde{\Theta}, \mathcal{M}, \Delta\}$ obeys the principle of ontic definiteness if and only if there exists some subset Θ of the ontic history space Ω from which $\tilde{\Omega}$ is derived, such that $\Theta \subset \tilde{\Theta}$ and all of the indefinite truth valuations in $\tilde{\Theta}$ can be derived from subsets of Θ .

In this way we can derive a three-valued theory from any ontic theory, simply by allowing for the truth values of some events to be unknown. In the card analogy, we now have a larger deck of cards, each one of which represents a truth valuation in $\tilde{\Theta}$. If we know the truth values of some events, we can again take all the cards that agree with them and throw the rest away. As before we can use the remaining cards to make predictions and draw inferences. In this case, there will be a unique card in our remaining hand, the “actual epistemic truth valuation”, on which all events with unknown truth values have truth value 1/2. But also, there will be many other cards representing other three-valued truth valuations which could become applicable when we come to know more. Amongst these, because of ontic definiteness, our hand will always include at least one of our original set of cards, the ontic histories.

¹²It is tempting to simply name this condition “realism”. This would chime somewhat with the usage in philosophy, especially the work of Dummett [40]. However, this word has long been used in discussions of quantum mechanics to mean something along the lines of “the opposite of operationalism”, which is certainly not intended here. Indeed, one of the main points to be made here is that the rejection of this condition need not dismay physicists who fall into the realist camp.

4.3 Einstein locality revisited

What becomes of our causal condition in this setting? If our original ontic theory $\{\Omega, \Sigma, \Theta, \mathcal{M}, \Delta\}$ obeys Einstein Locality, does this impose a condition on the three-valued theory $\{\tilde{\Omega}, \Sigma, \tilde{\Theta}, \mathcal{M}, \Delta\}$ derived from it? First let us consider the term “correlation” that we defined earlier for definite histories. There, two events A and B are correlated if and only if $\gamma(A) = \gamma(B) \forall \gamma \in \Theta$, which is equivalent to $A \cap \Theta = B \cap \Theta$. If this is true, what does it mean in terms of the three-valued truth valuations in the derived three-valued theory? Any set $X \subset \Theta$ is either a subset of $A \cap \Theta$, a subset of $A^c \cap \Theta$, or it intersects both. If $A \cap \Theta = B \cap \Theta$, in the first case we have $\tilde{\gamma}_X(A) = \tilde{\gamma}_X(B) = 1$, in the second case we have $\tilde{\gamma}_X(A) = \tilde{\gamma}_X(B) = 0$ and in the third case we have $\tilde{\gamma}_X(A) = \tilde{\gamma}_X(B) = 1/2$. So if A and B are correlated in the ontic theory, it follows that $\tilde{\gamma}(A) = \tilde{\gamma}(B) \forall \tilde{\gamma} \in \tilde{\Theta}$ in the derived three-valued theory. It therefore makes sense to adopt this as the definition of correlation for three-valued theories. Note that the definition for the three-valued case seems as immediately reasonable, on its own terms, as the original definition (indeed, it has the same form).

Now consider the following:

Einstein locality (three-valued formulation): Consider an indefinite theory defined by $\{\tilde{\Omega}, \Sigma, \tilde{\Theta}, \mathcal{M}, \Delta\}$, and regions \mathcal{X} , $\mathcal{R} = J^-(\mathcal{X})$ and $\mathcal{P} = J^-(\mathcal{X}) \setminus \mathcal{X}$. The theory is *Einstein local* if the following condition holds for all regions \mathcal{X} :

$$\forall A \in \Sigma|_{\mathcal{R}}, \exists B \in \Sigma|_{\mathcal{P}} \text{ such that } \tilde{\gamma}(A) = \tilde{\gamma}(B) \forall \tilde{\gamma} \in \tilde{\Theta}.$$

This is almost identical to the second formulation of Einstein locality given in section 2.1. It is not hard to see that this condition follows if the theory $\{\tilde{\Omega}, \Sigma, \tilde{\Theta}, \mathcal{M}, \Delta\}$ is derived from an ontic theory obeying EL. The second formulation of EL states that every event A that occurs in \mathcal{R} is correlated to some event B that occurs in \mathcal{P} . If two events are correlated in this sense in the ontic theory, it implies that $\gamma(A) = \gamma(B)$ for all $\gamma \in \Theta$ as well, as explained above. This shows that the three-valued formulation of EL follows from the more standard definite ones. It is also important to note that this formulation reduces to the definite formulation when only the original ontic truth valuations are considered.

This new formulation of EL has appeal independent of its derivation from the definite case. We have discovered that the meaning of correlation in the three-valued case should be that $\tilde{\gamma}(A) = \tilde{\gamma}(B)$ for all $\tilde{\gamma} \in \tilde{\Theta}$. If the truth values in question are in $\{0, 1\}$ this reduces to the normal definition, and when $\tilde{\gamma}(A) = \tilde{\gamma}(B) = 1/2$ the definition is also reasonable: if A is determining B (or if they are both determined by some other event) it would make little sense if one held a definite value and the other did not. With this in mind, the new version of EL simply states that every event is determined by an event in its past.

We also need to make sure that the ban on superluminal signalling follows from EL, as before. This is not difficult. Again, we consider a signal event A associated to a region \mathcal{A} , which has to be a “free setting”, and the output event B associated to a region \mathcal{B} which is spacelike to \mathcal{A} . In this case, we also have to note that both A and B must hold definite values; it would not be a signal if the message was unknown

to one or both parties. Freedom of the settings will still mean that we do not want any event C in $J^-(\mathcal{A}) \setminus \mathcal{A}$ (at least any such event that we need to include in our model) to be even weakly correlated to A (or A^c). As in the case of correlation, weak correlation also needs a counterpart for the three-valued theory.

In the ontic theory, weak correlation between C and A means that $C \cap \Theta \subset A \cap \Theta$. Given that the setting A has a definite value, what possible three-truth values can C and A take in a three-valued truth valuation $\tilde{\gamma}_X(\cdot)$? If X lies completely outside A (*i.e.* $X \subset A^c$), so that $\tilde{\gamma}_X(A) = 0$, then $\tilde{\gamma}_X(C) = 0$ also. If on the other hand $X \subset A$, then X can still be completely inside or outside C , or neither, so in this case $\tilde{\gamma}_X(C)$ can hold any of the three truth values. We find that, when A is definite, weak correlation between C and A should mean that $\tilde{\gamma}_X(C) = 0$ whenever $\tilde{\gamma}_X(A) = 0$ for all $\tilde{\gamma} \in \tilde{\Theta}$. Again, the definition is reasonable in its own terms.

Again, we exempt A from Einstein locality because we have deliberately left out its causal antecedents from our model. With three-valued truth valuations, the argument works exactly as before: signalling would mean that B is weakly correlated to A , whereas EL implies that there is an event C associated to the past of \mathcal{A} such that $\tilde{\gamma}(B) = \tilde{\gamma}(C) \forall \tilde{\gamma} \in \tilde{\Theta}$. But in that case C must also be weakly correlated to A , which violates our freedom of settings assumption. So superluminal signalling is, unsurprisingly, banned in any three-valued theory obeying EL in the above sense. For the following argument, it is crucial to note that this argument for a ban on superluminal signalling refers only to the three-valued version of EL and the three-valued meaning of correlation and weak correlation; we did not have to mention the ontic level at all.

So far, this discussion should have been quite comfortable even for those who are most skeptical of altering logic. After all, this discussion has been nothing more than a rewrite of the standard situation to formalise the case in which some truth values are unknown. Of course, this rewrite alone cannot save us from Bell's theorem¹³. Thus, adding a third truth value does not, on its own, go beyond the standard picture. The real break with the standard picture is yet to come. In order to maintain a satisfying causal principle even when faced with Bell's theorem, we must reject the principle of ontic definiteness.

4.4 The break with the standard picture

What does it mean to reject ontic definiteness? It means that our theory is an three-valued theory $\{\tilde{\Omega}, \Sigma, \tilde{\Theta}, \mathcal{M}, \Delta\}$ as before, but that the truth valuations in $\tilde{\Theta}$ need no longer contain the two-valued truth valuations corresponding to the histories in Θ from which they were derived. In other words, we no longer maintain that, although we do not know all truth values, nonetheless “in reality” all events have a definite truth value that we are simply ignorant of. We accept that, in some

¹³To prove Bell's theorem in this context we would still need to give some idea of how to put probabilities on our histories. Assuming ontic definiteness, it suffices to say that we would put probabilities only on the ontic histories, or alternatively that if we did entertain probabilities for events to be intermediate, these, and all their logical combinations, would be entirely independent of the probabilities on the ontic histories. Bell's theorem can then be derived as usual.

circumstances, there may be no meaning to hypothesising about the truth or falsity of certain propositions. To remind us of this change in view, we will rename the intermediate truth value “indefinite”. However, in order to avoid logical chaos, we have retained the other rules given above for the three-valued truth valuations. In particular, in order to be an acceptable three-valued truth valuation, it must be possible to derive the truth valuation from members of some standard history space Ω in the way described in section 4.2. The point can be illustrated with a simple example.

Let us start by assuming ontic definiteness. In this simple theory there are four allowed ontic histories. Event A corresponds to the proposition “the box is open”, while event B corresponds to the proposition “the ball is in the box”. The four possible histories cover every combination of truth and falsity here. Thus, two of the histories, $A^c \cap B$ and $A^c \cap B^c$, correspond to “the box is closed and the ball is inside”, and “the box is closed and the ball is not inside”. Finding the derived epistemic theory to go along with this ontic theory gives us histories signifying things like “the box is closed and it is intermediate whether or not the ball is inside” and so on.

Now let us reject ontic definiteness. We are now free to imagine a theory in which it is in principle impossible to know whether the ball is in the box when it is closed. If we want such a theory, we can consider just three histories: “the box is closed and it is indefinite whether or not the ball is inside”, “the box is open and the ball is inside”, and “the box is open and the ball is not inside”. These can be called the “underivable truth valuations”, or more simply the “basic truth valuations”, which now play a role similar to that which the ontic histories played before¹⁴: all other histories in the theory are derived from them in the way described in section 4.2. But now, if the box is closed, no allowed truth valuation gives a definite truth value to the event B .

This might bring Shrödinger’s cat to mind. However, rejecting ontic definiteness does not imply that we *must* treat *all* boxes in the way described here, but merely that we are liberty to describe *some* cases like this.

The most important thing to note here is that *our accounts of Einstein Locality and the ban on superluminal signalling still make perfect sense even if we drop ontic definiteness*. Thus we see that objections (1) and (2) against altering logic given in section 4.1 are unfounded in this case. By rejecting ontic definiteness we do no damage at all to the foundations on which the principle of causality is built. Instead we have, so to speak, pulled away the table cloth while leaving the dinner set untouched.

¹⁴There seem to be two ways of thinking here. We can think of the basic truth valuations ontologically, in which case we might be tempted to say that the intermediate events in the basic truth valuations are “indefinite in reality”, *i.e.* that the truth value $1/2$ as it appears in the basic truth valuations is *ontologically distinct* from 0 and 1. On the other hand, we can think of them epistemically, in which case we reject the idea that there is anything like an ontic history at all, and give up the idea of the basic truth valuations as possible “mirrors of reality” (something similar to Dummett’s “antirealism” position), in which case we may prefer to say “unknown” rather than “indefinite”. The distinction can be made – or considered meaningless – according to the taste of the reader as it has no effect on any of the following arguments. The important thing is that the “card game” described previously works just as well with or without ontic definiteness.

Is there any reason to preserve ontic definiteness? To argue that it cannot be sacrificed, one would have to come up with a reason why it is a useful or necessary principle. However, ideas very similar to removing ontic definiteness (such as rejecting the law of the excluded middle in formal logic) have often been considered by philosophers and mathematicians [40, 39]¹⁵. Besides this, it is commonly held that quantum mechanics is telling us that not all the properties of a system can be known even in principle; dropping ontic definiteness while preserving the logical structures presented above is a particular formalisation of this view.

This new picture is certainly different from more naïve versions of realism. But, it can be argued, it has all the attributes that those who call themselves realists (at least when it comes to quantum mechanics) would desire. Firstly indefinite logic does not imply operationalism. That is, it does not rule out the attribution of truth values to anything but operational events concerning preparations and measurements. We can still infer definite truth values about non-operational events by applying such an indefinite theory. So no special place need be given to the concepts of preparation or measurement. To the contrary, applying Einstein Locality will often *require* us to hypothesise events beyond the operational level to serve as explanations of correlations. Secondly, in this new picture, truth is still universal: it is not the case here that a proposition that “appears to be true” to an observer may in fact hold some different truth value in the greater scheme of things. There is only one universal truth value, 1, signifying “true”. Reliable observers should agree on all definite truth values in the most obvious sense. In these senses, indefinite logic provides all the objectivity that realists demand.

In addition, the new picture supplies something that previously seemed to be impossible: even while maintaining a satisfying causal principle, consistency with the results of quantum mechanics is possible.

4.5 The EPRB experiment

We now turn to the EPRB experiment to see how rejecting ontic definiteness allows us to evade Bell’s theorem. In our model, there will be sixteen basic histories in $\tilde{\Theta}$. We have two spacelike regions \mathcal{A} and \mathcal{B} . Region \mathcal{A} contains a setting event A_s and a result event A_r , and similarly for \mathcal{B} , we have B_s and B_r . For instance, in the standard EPRB scenario with spin-1/2 particles, A_s could correspond to the proposition “The Stern-Gerlach detector is set to angle θ rather than ϕ ” and A_r could mean “the outcome is spin up”. We also require that, within \mathcal{A} , A_s and A_r are associated to two subregions $\mathcal{A}_s \subset \mathcal{A}$ and $\mathcal{A}_r \subset \mathcal{A}$, such that \mathcal{A}_s is to the past of \mathcal{A}_r , and similarly for \mathcal{B} . This is done so that the setting counts as being to the past of the result.

In the mutual past $\mathcal{P} = J^-(\mathcal{A}) \cap J^-(\mathcal{B})$ we will include four more events, P_{A0} , P_{A1} , P_{B0} and P_{B1} . There are sixteen allowed truth valuations in $\tilde{\Theta}$ corresponding to the 16 possible definite truth valuations on the settings and results in regions \mathcal{A} and

¹⁵In their discussion of “schematic worlds”, Rescher and Branden come particularly close to the scheme considered here, especially if the truth-value $/P/$ defined in section 8 of [39] is compared to $\tilde{\gamma}(P)$ as used above.

\mathcal{B} . The truth values for the past events are decided as follows. For truth valuations in which setting A_s is valued true, *i.e.* when $\tilde{\gamma}(A_s) = 1$, P_{A1} holds the same truth value as A_r while P_{A0} is indefinite; when $\tilde{\gamma}(A_s) = 0$, P_{A0} holds the same truth value as A_r while P_{A1} is indefinite. The situation for the B variables is the same with B substituted for A . The truth values of these eight events are given for all of the sixteen histories in a table to be found in appendix D.2.

Specifying these truth values does not specify the truth values for all logical combinations of these events, and so to give the whole truth valuation we need to supply more. For every truth valuation in Θ , two of the events P_{A0} , P_{A1} , P_{B0} and P_{B1} have indefinite truth values (*e.g.* , when $\tilde{\gamma}(A_s) = \tilde{\gamma}(B_s) = 1$, $\tilde{\gamma}(P_{A0}) = \tilde{\gamma}(P_{B0}) = 1/2$) and so the truth values of some of their logical combinations are not given by specifying their own truth values. Several specifications can be given here which are compatible with our physical conditions. This issue is treated at length in appendix D.

The following argument establishes that Einstein locality condition holds for this model. Exempting setting events from consideration as usual, it remains to show that the outcome events A_r and B_r are determined by some past event. It is not hard to verify that $\tilde{\gamma}((A_s \cap P_{A1}) \cup (A_s^c \cap P_{A0})) = \tilde{\gamma}(A_r)$ for all of the 16 basic truth valuations, and similarly for the B -labelled events. Because the result events are definite, their logical combinations are determined by logical combinations of their causal antecedents as usual. It is also apparent that freedom of settings holds. No past event in the model is weakly correlated to A_s or B_s for all of the basic truth valuations. That is, none of the past events holds the truth value 0 whenever A_s holds the truth value 0, and similarly for B , as is easily seen in the table. It is also true that no (non-trivial) logical combination of the past events is weakly correlated to any (non-trivial) logical combination of the setting events, as shown in appendix D, where these issues are treated in more detail.

Putting a probability distribution on these indefinite truth valuations (whilst maintaining freedom of settings in the probabilistic sense¹⁶) it is possible to maximally violate the CHSH inequalities. This is clear from the fact that the model puts no restrictions at all on the probability distributions over the events in \mathcal{A} and \mathcal{B} apart from no-signalling. Thus we can have maximal correlations – correlations even stronger than those allowed by quantum mechanics can be achieved here – while maintaining a causal principle that bans superluminal influence.

5 Discussion

5.1 The “meta-level” objection

The meta-level objection referred to before, in this case, is that we have exploited the “causal loophole” or (more or less equivalently) allowed backward causation. After all, when $\tilde{\gamma}(P_{A1}) = 1/2$, the truth value of the setting event $\tilde{\gamma}(A_s)$ can only be 0.

¹⁶Here this should mean that the probability of any event in the past, given that it is definite, is independent of the probability of the settings. A similar treatment of the GHZ experiment [41, 42] can be given with no reference to probabilities. This will be left for future work.

Doesn't this mean that P_{A1} is determining A_s ? This is true, *if* one treats $\tilde{\gamma}$ not as a truth valuation but as a list of variables describing an ontic history, and goes on to use standard logic to talk about these 16 histories. More precisely, we can invent a new “meta-algebra of events” $\hat{\Sigma}$ in which the “meta-events” correspond to the truth valuations of events given by the histories $\tilde{\gamma}$. For instance, in the above example “ $\tilde{\gamma}(P_{A0}) = 1/2$ ” is a meta-event. We can introduce a standard, definite truth valuation on $\hat{\Sigma}$, which we call a “meta-history” $\hat{\gamma}$, corresponding to each indefinite truth valuation $\tilde{\gamma}$. For instance, for all of the 16 meta-histories in our example, $\hat{\gamma}(\tilde{\gamma}(P_{A0} = 1/2)) = 1$ if and only if $\tilde{\gamma}(P_{A0} = 1/2)$, and so on. This is what is meant by the “meta-level” – in this way it is always possible to find a corresponding definite theory for any indefinite theory. This meta-level objection can be leveled at any such change in logic¹⁷. On the meta-level, standard logic applies, and we can define theories just as before. So, as before, one of the assumptions of Bell's theorem mentioned above must be violated. How then could using such an altered logic bring anything new to the discussion of Bell's theorem?

Even looked at from this “meta-level” point of view, there is something to be said for the altered logic set out above. There is still a set of conditions here that bans superluminal signalling – the indefinite version of EL (which, it can be seen, is *stronger* than the definite version at the meta-level), the indefinite version of freedom of settings (which is *weaker* than the definite version at the meta-level) and the condition that the setting and outcome variables can have the values 0 or 1 but never 1/2. Together, these ban superluminal signalling. From this point of view what we have done is to strengthen Einstein locality in such a way that we can ban superluminal signalling. So even if we take this meta-level stance, the worst that can be said is that we are utilising the freedom of settings loophole. Still, a condition can indeed be stated which bans superluminal signalling and yet allows violations of Bell's inequalities, which is surely progress of the type we were looking for. Apart from this, though, motivation for this new “causal principle” is somewhat opaque from the “meta-level” point of view.

While this is certainly a possible stance to take, I will argue that it is unreasonable to insist on this “meta-level” viewpoint. There is no justification for the claim that the three-valued version of Einstein Locality is “not really” Einstein Locality, when it serves exactly the same function as the definite version and is derived on exactly the same grounds. The same holds for the three-valued version of freedom of settings. We have already seen that they are motivated in the same way as the standard, definite expressions, and hopefully they can be just as useful for the progress of physics. Furthermore, the “card game” described above is intuitive and standard for three-valued theories *with* ontic definiteness, and this game proceeds in exactly the same way when ontic definiteness is dropped; in the light of this there is no reason to retreat to a meta-level perspective in order to ensure that we can reason effectively. Additionally, for any definite events (events for which we know the truth value) we can reason exactly as usual. This provides a strong argument that the indefinite

¹⁷Or at least, any change in logic for which the truth of an event as determined by a reliable observer is universal.

statements are just as worthy of their names as their definite cousins, and that this way of speaking can be just as meaningful, and just as useful – in fact more so because we are no longer in contradiction with quantum mechanics¹⁸.

It is true that, when $\gamma(P_{A1}) = 1/2$, the setting event A_s can only be false. From the indefinite logic point of view there is nothing odd about that. Technically it does not violate our freedom of settings requirement because P_{A1} is not weakly correlated to A_s . To gain an intuition about this, we can imagine a theory that contains 3 basic indefinite histories, corresponding to the statements “the box is closed today and whether the ball was in it yesterday or not is indefinite”, “the box is open today and the ball was in it yesterday” and “the box is open today and the ball was not in it yesterday”. For this perspective there seems to be nothing strange in saying that an event that happened in the past (the ball being in the box yesterday) can be made definite or indefinite by an event that happens today (opening the box to check). In particular, it would be unreasonable to claim that the later event was determined by the earlier one (or *vice-versa*). All that is happened is that, if the box is open, this gives us grounds to attribute a definite truth value to whether the ball was in there yesterday, which are otherwise absent. Thus, it can be the case that one event (even a past event) is “made definite” by another without them being correlated. This is simply not what we mean by correlation, in this context.

5.2 The relation to counterfactual definiteness

The connection to the idea of counterfactual definiteness is now apparent. By dropping the assumption of ontic definiteness, we can revive this “loophole”. In the treatment of the EPRB experiment given above, it is indeed the case that no definite value is attributed to the results of the measurements that are not performed. This is therefore an example of lack of counterfactual definiteness. In this setting however, the lack of counterfactual definiteness is not vulnerable to Norsen’s criticism: he derives counterfactual definiteness from the assumption of local causality (note that he could just as easily have assumed the slightly stronger Einstein Locality to reach the same conclusion) and freedom of settings, but assuming ontic definiteness. Instead, we have dropped Norsen’s (and Bell’s) assumption of ontic definiteness, while preserving the indefinite versions of of Einstein locality and freedom of settings. This provides a way to evade Norsen’s argument that counterfactual definiteness is required by local causality. It should, again, be stressed that making this move does not commit us to operationalism, against which Norsen directs his argument [33].

5.3 Relation to other programmes

The idea that ontic definiteness can be dropped, allowing the preservation of a meaningful relativistic principle of causality without contradicting quantum mechanics,

¹⁸An argument for a similar claim is made at length in [39]. There, it is argued that Rescher and Brandom’s inconsistent semantics need not be considered as merely derivative from the standard one, but could just as easily be considered the basic case from which the standard one is derived, a stance the authors call “the parity thesis” (see in particular section 13 of [39]).

should be of interest generally in inquiries into the foundations of quantum mechanics. It is also of specific interest to more than one programme for interpreting quantum mechanics. On its own, dropping ontic definiteness does not provide an interpretation, but only a new way to evade Bell's theorem. Nevertheless, as has been noted, one of the prime decisions to be made when interpreting quantum mechanics (whether in an operationalist way or not) is what assumption of Bell's theorem to drop. Using the newly uncovered assumption to build a full interpretation is surely of interest.

The most obviously related programme is the Ψ -epistemic approach. Here, the quantum state is viewed, not as reflective of an element of reality (*i.e.* not ontologically), but instead as expressing our state of knowledge about the system in question. Spekkens has proposed a toy model relying on these ideas that qualitatively reproduces many interesting features of quantum mechanics [4]. He introduces a principle under which, in a particular well-defined sense, the truth values of only half of all events can ever be known at once, thus limiting the possible epistemic states to always be “coarser grained” than the ontic states. This is moving very near to the direction this article has taken. However, although the possible epistemic states are limited in the toy model, it is still assumed that ontic states exist, from which the epistemic states are derived as above in section 4.2. As we have seen, jettisoning this assumption allows us to evade Bell's theorem. It would be most interesting to see if Spekkens's toy model can be improved by removing the assumption of ontic definiteness, allowing it to model *all* interesting features of quantum mechanics, including the stronger-than-classical correlations. For instance, perhaps the model could be altered to produce “stabiliser-group quantum theory” (a reduced version of standard quantum theory), which has been found to be very similar to the toy model [43]. If this could be achieved, the new toy model would indeed be able to do better than the original one in modelling the strong correlations found in GHZ states [41, 42].

Connections can also be found to the anhomomorphic logic program [24, 25, 26]. In this approach to quantum mechanics, as here, the algebra of events is Boolean as in standard logic, and their truth values are still in $\{0, 1\}$, but the truth valuation need not be homomorphic (so, for example, if A is false and B is false it does not necessarily follow that $A \cup B$ is false). There are many possible ways to weaken the rules of classical logic after abandoning homomorphic truth valuations. Interestingly, the one that has turned out to be most useful for quantum mechanics, the “multiplicative scheme”, bears strong similarities to the ideas set out above. That is, there is a one-to-one map between truth valuations and subsets of the history space Ω , as above. In the truth valuation corresponding to subset X , the event A is true if and only if $A \supset X$. This is also true for the derived three-valued truth valuations described above, as noted in section 4.2. The difference is that, in anhomomorphic logic there are only two truth values, and so all other events are designated false (whereas, in the derived three-valued truth valuations, only events A such that $A^c \supset X$ are false, and the rest are intermediate). We see here that for every truth valuation in anhomomorphic logic, it is easy to substitute an indefinite truth valuation. The events given indefinite truth values are exactly those events A for which, in the anhomomorphic “co-event”,

A and A^c are both deemed false¹⁹. Methods have already been proposed to derive the set of possible anhomomorphic truth valuations from quantum mechanical theories. This supplies a “ready-made” interpretation of quantum mechanics with indefinite histories and no ontic definiteness. However, one would still have to prove in each case that the theory obeyed Einstein locality (or the weaker relativistic causal principle hinted at near the end of section 2.1) in this interpretation. As yet, the theory contains no general rule that explicitly ensures this.

It is also natural to wonder about the connection to the topos-based approach to quantum interpretation pioneered by Butterfield, Döring and Isham [22, 23]. This approach also alters logic, but similarities do not go much further. In this approach, what an observer would normally call true should properly be considered only “true to them”, or true only in their particular context. This is in contrast to the case discussed above where there is one truth value meaning “true”, 1, which is universal. This is the main reason why it is difficult simply to transpose the discussion of causal principles given here to the case of the topos-based approach. The discussion of whether a satisfying causal principle can be maintained in that interpretation may, for this reason, have more similarities to the discussion for many-worlds approaches.

6 Conclusion

This paper began by noting that Bell’s theorem seems to prevent us from maintaining a satisfying relativistic principle of causation. When interpreting quantum mechanics, we seem to be faced with a choice between maintaining the principle of common cause, or relativistic causal structure. Thus the road chosen here depends on whether or not we see causality as a necessary part of any satisfying theory.

In this article, it has been shown that there is a hidden assumption that forces us to make this choice. Ontic definiteness is the principle that, whether we can (even in principle) have grounds to know them or not, every proposition about physical systems must have a definite truth value, “true” or “false”. We have seen that this principle in no way underpins either the causal principle of Einstein Locality, or the assumption of freedom of settings used in Bell’s theorem. However, it does underpin Bell’s derivation of a contradiction with the predictions of quantum mechanics, starting with these principles. We are left with a choice: we can either abandon the demand for a satisfying relativistic causal principle, or we can abandon the principle of ontic definiteness. Abandoning ontic definiteness seems to do no damage to the demands made by physicists who defend realism: we are still allowed, and moreover required, to make inferences about facts beyond the operational level. On the other hand, those who reject approaches like deBroglie-Bohm pilot wave theory on the grounds that they contradict locality should find nothing to object to either.

There is a strong case to be made that, if we have to do without one, maintaining ontic definiteness is less desirable than maintaining a satisfying causal principle. Arguments have been given above that no other way of evading Bell’s theorem allows a

¹⁹This line of thought also suggests that the above evasion of Bell’s theorem could be re-written in the anhomomorphic logic language.

satisfying causal principle to be maintained. When interpreting quantum mechanics, therefore, dropping ontic definiteness is not only a new option, but is arguably the most attractive one available.

This raises the question of what such an interpretation would look like. Above, it was suggested that simple models in the spirit of Spekken’s toy model could be constructed to gain a better understanding of this. One of the major questions here concerns the strength of correlations in quantum theory. Without ontic definiteness, Einstein locality allows the strongest non-signalling correlations that are logically possible, while quantum mechanics places restrictions on the strength of correlations. It is very interesting to ask whether this can be explained as the result of any simple set of physical principles. These issues are of current interest for those studying foundational issues in quantum information, for example. The same question arises for this new programme. If Einstein locality alone does not imply the quantum restrictions on correlations, will it do the job when supplemented by other physical principles, and if so, what are they? Hopefully, toy models can shed more light on this issue.

Other interesting extensions include the application to quantum gravity. It is arguable that a proper understanding of quantum mechanics is a prerequisite for this. But even so, it is interesting to ask at this stage what the principle of Einstein Locality would look like if we removed the assumption of a fixed background spacetime. With general covariance in mind, this may require answers to some deep questions, along the lines of “what is a region?”. Perhaps addressing the issue in the conceptually simpler context of causal sets will lead to progress here.

There are many other avenues for future work. One is to study Mermin’s formulation of the GHZ experiment [42] in the light of the ontic definiteness assumption. Beyond this, it would be of use to study indefinite logic independently of the derivation from standard logic, and to consider the role of probabilities more carefully.

We have seen that principles of causation, and the physicist’s notion of realism, are closely intertwined concepts. By further studying the consequences of rejecting ontic definiteness, it should be possible to make further progress in the task of restoring a satisfying picture of the microworld to our physical theories.

Acknowledgements

The author is grateful to Jeremy Butterfield, Adam Caulton, Fay Dowker, Adrian Kent and Rafael Sorkin for useful discussions, and to Rob Spekkens for pointing out the work of Travis Norsen on counterfactual definiteness.

A Causal Principles: what are they good for?

A.1 Defending the causal principle

In section 2 it was asserted that finding a satisfying causal principle that is compatible with quantum theory may help us to make scientific progress. A few comments by

way of clarification and justification of this statement are given in this appendix.

Causation furnishes us with an important mode of explanation. Requiring causal explanation often means requiring the hypothesis of new events (in the sense in which the word is used above) to serve as common causes for otherwise unexplained correlations. This is part of the story of discovering new substances and processes. As a simple example, consider a case in which the movement of one tree is correlated to the movement of another tree. Let us assume that we have good reason to think that the movement of one tree is not in itself causing the movement of the other. In this case, it would be strange to explain the simultaneous movement of two trees by stating that it is in the nature of trees to move simultaneously, and leave it at that. Instead, the PCC requires us to posit a new effect that furnishes the common causes for the simultaneous movement of trees. In this way we enrich the subjects of our physical inquiry from the initially apparent events (which all have to do with the properties of trees) to new events (having to do with the wind).

Indeed, having gained confidence in our wind theory, we make no fundamental division between the events we originally categorised as “apparent” and those which we inferred from them using the causal principle. The boundaries of what is considered operationally “apparent” and what is not can also change if we take the causal principle seriously: for example the electromagnetic field, at first taken as inferred through its effects on matter, is in modern theories no more or less deserving of a special “apparent” status than matter fields. It is interesting to note that taking a strictly operational point of view (though it might be useful as a means of shedding unnecessary assumptions present in previous theories) prevents this kind of advance, by fixing the divide between the events that are considered apparent and the rest. Thus, the causal principle is closely intertwined with the realistic (*i.e.* non-operational) world-view. Moreover, it is interesting to ask: is there any use in insisting on a non-operational point of view *without* also insisting on a satisfying causal principle? In quantum mechanics we may well be able to solve the measurement problem by denying the PCC. But in so doing, we give up the possibility of enriching the subjects of our physical inquiry by using a causal principle.

A.2 Banning superluminal signalling is not enough

From an operational standpoint, it is tempting to ask: if we want a causal principle that bans superluminal signalling, why not simply take the ban on superluminal signalling itself to be the causal principle? However, there is a good reason to make a distinction between the two ideas. In the example of the previous section, we might shake one of the trees, and observe that distant trees are not caused to move. This certainly adds to the evidence that there is no direct causal link between the movement of the two trees. But it does not do anything to generate an *explanation* of why the trees’ movement is correlated despite this, leading to our discovery of the effect of wind. A true causal principle, like the PCC, can do this for us. As argued in the previous section, one of the points of reinstating a causal principle is to be able to talk about events that are not directly observable (or not at first considered to be), and so it defeats the object to retreat to an operational point of view .

There is another related reason not to accept the ban on superluminal signalling as a satisfying causal principle: doing so would confuse superluminal *influence* with superluminal *signalling*. For example, imagine that we often observed a bright light being emitted from every point on a spacelike path through spacetime. When trying to design a physical theory that explained this effect, it would be very unnatural to aim for a theory in which there was no superluminal influence. But it could still be the case that these effects could not be controlled and used to signal. A reasonable theory here could maintain a ban on superluminal signalling, but not a ban on superluminal influence in general. It is superluminal influence in the more general sense that is banned by a satisfying relativistic causal principle.

A.3 The challenge of causal antifundamentalism

Some recent arguments have been made that the idea of a causal principle should, at this stage in the development of science, be rejected as a universal principle for science. This is not the converse of the “modest assertion” made at the beginning of section 2, and so need not concern us too much, but the claims still deserve some discussion in this context. Butterfield has also discussed Norton’s stance in the context of relativistic causality and quantum mechanics in [44].

Arthur Fine questions whether correlations need explanation [45]. Essentially, he is advocating dropping the PCC as a response to the problems of quantum mechanics. Hopefully the arguments of the previous subsection are sufficient to give question to this. John Norton gives a more historically-based criticism of causation as a fundamental principle of science [31]. He argues:

Mature sciences, I maintain, are adequate to account for their realms without need of supplement by causal notions and principles. The latter belong to earlier efforts to understand our natural world, or to simplified reformulations of our mature theories, intended to trade precision for intelligibility. In this sense I will characterize causal notions as belonging to a kind of folk science, a crude and poorly grounded imitation of more developed sciences.

One part of Norton’s argument is that our idea of causation has had to be redefined after major developments in physics, leading one to suspect that there exists no general form for what he calls “causal notions” at all. He also argues that some examples even in Newtonian mechanics show violations of causal principles. Finally he brings up the very real problem of Bell’s theorem. If Norton’s position, which he dubs “causal antifundamentalism”, is correct, we can conclude that theories should not be considered inadequate because they fail to obey a causal principle (like a version of the PCC), since causal principles are simply folk-science and may have already become obsolete.

Norton’s position deserves serious consideration. In particular it is hard to argue that the casual principle must be defined once and for all, and preserved through

every possible development in science. Nevertheless, I will argue that it is worthwhile to preserve a well-defined causal principle for the time being.

To the first point, the fact that “causal notions” have had to be refined, and that, at some points in history, our best understanding of them has been considered problematic, need not suggest their eventual dissolution. This has not, for example, been the case for “gravitational notions”. Neither are the pathological examples given from Newtonian mechanics very worrying for those who seek to reinstate the idea of a causal principle in modern physics. To start with the examples are based on highly unphysical examples (in one example, a perfect ball must be positioned at an exact point on an object with an infinitely precise profile). They can, no doubt, easily be removed by appending reasonable and near-trivial restrictions to Newtonian mechanics, rendering the theory (non-relativistically, deterministically) causal. More importantly for practical purposes (although this does not contradict Norton), the use of causal arguments may yet have been important to the development of various theories despite their failure in pathological cases. The remaining argument is that we have managed to make huge advances in quantum mechanics, a theory in which the causal principle is still mysterious. This is certainly an important point to consider. But it proves neither that there is no satisfying causal principle to be found here, nor that it is not advisable to find and use it.

Norton’s argument that modern theories do not *need* to be supplemented by causal principles in order to function does not affect my argument that causal principles can be a useful tool for understanding them, in order (amongst other things) to develop yet more powerful theories. One explanation of correlations between events is that our theory allows (or requires) such correlations. But this presupposes that such a theory exists. This “static” view of theory ignores a crucial role of causation as a means of developing theories, as the tree example elucidates.

Finally, it is interesting to wonder why, if it is possible, anyone would object to reinstating a satisfying causal principle in quantum mechanics. Before Bell’s theorem was known, this was a possibility that had yet to be checked. Objecting even to attempts to reinstate the causal principle in these circumstances seems quite perverse. Bell himself used the analogy of “the glove left at home” to explain this:

If I find that I have bought one glove, and that it is right-handed, then I predict confidently that the one still at home will be seen to be left-handed. But suppose we had been told, on good authority, that gloves are neither right- or left-handed when not looked at. Then that, by looking at one, we could predetermine the result of looking the other, in some remote place, would be remarkable. Finding that this is so in practice, we would very soon invent the idea that gloves are already one thing or the other even when not looked at. And we would begin to doubt the authorities that had assured us otherwise.

[3] Indeed, this is the substance of the EPR debate, as Bell notes. Elsewhere he is quoted as saying of the EPRB experiment:

For me, it is so reasonable to assume that the photons in those experiments carry with them programs, which have been correlated in advance, telling

them how to behave. This is so rational that I think that when Einstein saw that, and the others refused to see it, *he* [not Bohr] was the rational man... so for me, it is a pity that Einstein's idea doesn't work. The reasonable thing just doesn't work.

[46] Here, the view set out in the present paper is that Bell is, although not in so many words, mourning the loss of a satisfying relativistic causal principle to modern physics due to its apparent incompatibility with the predictions of quantum mechanics. Above, it is shown that Bell's impasse can be avoided by dropping the assumption of ontic definiteness. It is true that dropping this principle is a greater leap than Einstein initially foresaw when he proposed that quantum mechanics should be causally "completed" while respecting relativistic "locality". But on the other hand, since a road towards Einstein's goal is now open, it seems unreasonable to advocate against pursuing it.

B Comparing causal principles

B.1 The two formulations of Einstein locality

In section 2.1 it was claimed that the two definitions of Einstein Locality that were given were equivalent. This claim is proved in this appendix.

First let us assume the first formulation and derive the second. For every history $\gamma_{\mathcal{P}}^* \in \Omega|_{\mathcal{P}}$ of events in the past region \mathcal{P} , there exists a history $\gamma_{\mathcal{R}}^* \in \Omega|_{\mathcal{R}}$ for \mathcal{R} such that for all $\gamma \in \Theta$, if $\gamma|_{\mathcal{P}} = \gamma_{\mathcal{P}}^*$ then $\gamma|_{\mathcal{R}} = \gamma_{\mathcal{R}}^*$. This implies the following:

$$\forall \gamma \in \Theta, \gamma' \in \Theta, A \in \Sigma|_{\mathcal{R}}, \gamma|_{\mathcal{P}} = \gamma'|_{\mathcal{P}} \Rightarrow \gamma(A) = \gamma'(A). \quad (1)$$

In words, all allowed histories that match on \mathcal{P} give the same truth value to all events associated to \mathcal{R} . This allows us to construct an event associated to \mathcal{P} that shares the same truth value as A for all allowed histories:

$$B = \{\gamma \in \Omega : \gamma|_{\mathcal{P}} = \gamma'|_{\mathcal{P}} \Rightarrow \gamma'(A) = 1 \forall \gamma' \in \Theta\}. \quad (2)$$

For any $\gamma_{\mathcal{P}}^* \in \Omega|_{\mathcal{P}}$, the event $\{\gamma \in \Omega : \gamma|_{\mathcal{P}} = \gamma_{\mathcal{P}}^*\}$ is associated to \mathcal{P} by the definition of $\Omega|_{\mathcal{P}}$, and the event B is associated to \mathcal{P} by virtue of being the union of a set of such events. It remains to show that the event B is correlated to A . Consider an allowed history $\gamma \in \Theta$ such that $\gamma(A) = 1$. Any such history will fulfil the condition in (2) due to (1), and so it will be in B , giving $\gamma(B) = 1$ in this case. Any $\gamma \in \Theta$ such that $\gamma(A) = 0$ will fail the condition of (2) for the same reason, and will not be in B , giving $\gamma(B) = 0$. Thus $\gamma(A) = \gamma(B)$ for all $\gamma \in \Theta$, or in other words, they are correlated. For an arbitrary event $A \in \Sigma|_{\mathcal{R}}$, we have found an event $B \in \Sigma|_{\mathcal{P}}$ that is correlated to A , and thus satisfied the second definition of Einstein locality.

Now let us consider the converse case. We start with the assumption that $\forall A \in \Sigma|_{\mathcal{R}}, \exists B \in \Sigma|_{\mathcal{P}}$ such that $\gamma(A) = \gamma(B) \forall \gamma \in \Theta$. Consider an event $A = \{\gamma \in \Omega : \gamma|_{\mathcal{R}} = \gamma_{\mathcal{R}}^*\}$ for some $\gamma_{\mathcal{R}}^* \in \Omega|_{\mathcal{R}}$. This event is associated to \mathcal{R} . Note that for all histories $\gamma \in \Omega$, $\gamma|_{\mathcal{R}} = \gamma_{\mathcal{R}}^*$ if and only if $\gamma(A) = 1$. By the second formulation of EL,

there is an event $B \in \Sigma|_{\mathcal{P}}$ correlated to A . In other words, for all histories $\gamma \in \Theta$ if $\gamma(B) = 1$ then $\gamma(A) = 1$, and so $\gamma|_{\mathcal{R}} = \gamma_{\mathcal{R}}^*$. The history restricted to \mathcal{P} , $\gamma|_{\mathcal{P}}$ fixes the truth value of B . Also, for every history $\gamma \in \Omega$ there exists some $\gamma_{\mathcal{R}}^*$ such that B is true. Thus, for every restricted history $\gamma_{\mathcal{P}}^* \in \Omega|_{\mathcal{P}}$, there exists a history $\gamma_{\mathcal{R}}^* \in \Omega|_{\mathcal{R}}$ for \mathcal{R} such that for all $\gamma \in \Theta$, if $\gamma|_{\mathcal{P}} = \gamma_{\mathcal{P}}^*$ then $\gamma|_{\mathcal{R}} = \gamma_{\mathcal{R}}^*$. We have recovered the first formulation of Einstein locality.

B.2 The two formulations of the Nonprobabilistic PCC

In 2.1 two formulations of a non-probabilistic principle of common cause are given, called NPPCm and NPCCj. These formulations are claimed to be equivalent. These claims are proved in this section.

The context is a nonprobabilistic theory defined by $\{\Omega, \Sigma, \Theta, \mathcal{M}, \Delta\}$, two spacelike regions \mathcal{A} and \mathcal{B} , their “joint past” $\mathcal{P}_j = J^-(\mathcal{A}) \cup J^-(\mathcal{B}) \setminus (\mathcal{A} \cup \mathcal{B})$ as shown in figure 2 and their “mutual past” $\mathcal{P}_m = J^-(\mathcal{A}) \cap J^-(\mathcal{B})$. The theory obeys the NPCCj if the following principle holds: for any pair of events $A \in \Sigma|_{\mathcal{A}}$ and $B \in \Sigma|_{\mathcal{B}}$, if A and B are correlated, then there exists a “common cause” event $C \in \Sigma|_{\mathcal{P}_j}$ that is correlated to both A and B . The definition of NPCCm is the same with \mathcal{P}_j replaced by \mathcal{P}_m .

First, it is easy to see that NPPCm implies and NPCCj. NPPCm requires that all pairs of correlated events A and B as described above have common cause event in the mutual past $\mathcal{P}_j = J^-(\mathcal{A}) \cap J^-(\mathcal{B})$. This is a subset of the joint past $\mathcal{P}_m = J^-(\mathcal{A}) \cup J^-(\mathcal{B}) \setminus (\mathcal{A} \cup \mathcal{B})$. Thus, if there is a common cause event in the mutual past then there is also one in the mutual past, and NPCCm is also satisfied.

Conversely, consider the regions $\mathcal{Y} = J^-(\mathcal{A}) \setminus J^-(\mathcal{B})$ and $\mathcal{Z} = J^-(\mathcal{B}) \setminus J^-(\mathcal{A})$. Because $\mathcal{A} \subset \mathcal{Y}$ and $\mathcal{B} \subset \mathcal{Z}$, $\Sigma|_{\mathcal{A}}$ is a subalgebra of $\Sigma|_{\mathcal{Y}}$ and $\Sigma|_{\mathcal{B}}$ is a subalgebra of $\Sigma|_{\mathcal{Z}}$. Thus $A \in \Sigma|_{\mathcal{Y}}$ and $B \in \Sigma|_{\mathcal{Z}}$. If the joint past version of the NPCC holds for all pairs of spacelike regions, then it must apply to \mathcal{Y} and \mathcal{Z} . The joint past of \mathcal{Y} and \mathcal{Z} is $\mathcal{P}_m = J^-(\mathcal{A}) \cap J^-(\mathcal{B})$. Applying the joint past version of the NPCC we see that, if A and B are correlated, then there exists an event $C \in \Sigma|_{\mathcal{P}_m}$ that is correlated to both B and C . Thus the mutual past version of the NPCC also holds. This proves that NPCCm follows from NPCCj, and as a corollary that NPCCm and NPCCj are equivalent. This result, and the proofs, are similar to those already known for the probabilistic case [29].

This result may seem questionable on first sight. Why can’t the entire chain of causal antecedents of A lie in $J^-(\mathcal{A}) \setminus \mathcal{A}$ but outside of $J^-(\mathcal{B})$, and similarly for B , so that there is no common cause in the mutual past? Yet we have proved that the NPCCm principles follows from NPCCj, which is justified on general grounds. Could there be something wrong with our assumptions, therefore? But we should consider what it means for the entire chain of causal antecedents of A to lie outside of $J^-(\mathcal{B})$. If there is an initial cosmological spatial hypersurface beyond which no past exists, this would imply that there are two events on that surface (one the causal antecedent of A , and the other of B), at spacelike positions, that are correlated to each other. NPCCj disallows these “initial correlations” as they have no common cause (similarly “correlations that propagate in from past infinity” are banned). This point is discussed in [29] for the probabilistic case.

B.3 Einstein locality and screening off

It was also claimed in section 2.1 that Einstein locality implied a certain stochastic relativistic causality condition when a probability distribution was put on Θ . This is now shown. Along similar lines, Bell comments that a deterministic version of causality follows from his local causality when probabilities are restricted to be 1 and 0 [3].

Consider two spacelike spacetime regions \mathcal{A} and \mathcal{B} , their union $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$, their joint past region $\mathcal{R} = J^-(\mathcal{A} \cup \mathcal{B})$ and their exclusive past $\mathcal{P} = \mathcal{R} \setminus (\mathcal{A} \cup \mathcal{B})$, as shown in figure 2. According to the first formulation of Einstein locality, for every history $\gamma_{\mathcal{P}}^* \in \Omega|_{\mathcal{P}}$ of events in the past region \mathcal{P} , there exists a history $\gamma_{\mathcal{R}}^* \in \Omega|_{\mathcal{R}}$ for \mathcal{R} such that for all $\gamma \in \Theta$, if $\gamma|_{\mathcal{P}} = \gamma_{\mathcal{P}}^*$ then $\gamma|_{\mathcal{R}} = \gamma_{\mathcal{R}}^*$. Since \mathcal{C} is a subregion of \mathcal{R} , it also holds that for every history $\gamma_{\mathcal{P}}^* \in \Omega|_{\mathcal{P}}$ of events in the past region \mathcal{P} , there exists a history $\gamma_{\mathcal{C}}^* \in \Omega|_{\mathcal{C}}$ such that for all $\gamma \in \Theta$, if $\gamma|_{\mathcal{P}} = \gamma_{\mathcal{P}}^*$ then $\gamma|_{\mathcal{C}} = \gamma_{\mathcal{C}}^*$. Similar statements also hold for the regions \mathcal{A} and \mathcal{B} .

Consider a probability distribution μ on Ω with support in Θ , and assume Einstein Locality. Consider also the event $P = \{\gamma \in \Omega : \gamma|_{\mathcal{P}} = \gamma_{\mathcal{P}}^*\}$, which we can call a “full specification” of \mathcal{P} . Conditioning on P is equivalent to fixing the history restricted to \mathcal{P} to be $\gamma_{\mathcal{P}}^*$. According to the first formulation of EL, this also fixes the history on \mathcal{C} to some value $\gamma_{\mathcal{C}}^*$, and hence fixes the the history in \mathcal{A} and \mathcal{B} , making all probabilities conditioned on P either 0 or 1. This means that, for all events C associated with \mathcal{C} , $\mu(C|P) = 0$ if $\gamma_{\mathcal{C}}^*(C) = 0$ and $\mu(C|P) = 1$ if $\gamma_{\mathcal{C}}^*(C) = 1$. For any pair of events A associated to \mathcal{A} and B associated to \mathcal{B} , the fact that $\gamma_{\mathcal{C}}^*$ is a homomorphism then implies that $\mu(A|P)\mu(B|P) = \mu(A \cap B|P)$. In other words, we have shown that all pairs of events A and B (associated to regions \mathcal{A} and \mathcal{B} respectively) are statistically uncorrelated when conditioned on any full specification of events in their joint past region \mathcal{P} . This “screening off” condition is called SO2 in [29]. This SO2 condition has been shown to be equivalent to the similar condition (called SO1) in which the “mutual past” $J^-(\mathcal{A} \cap \mathcal{B})$ is substituted for the joint past in the definition [29], which is essentially the same as one of Bell’s formulations of local causality (given in “The theory of Local Beables” [2]).

C Comparison with Reichenbach’s logic

In several important senses, the logic introduced in section 4.2 has a precursor in Reichenbach’s three-valued logic [21]. Reichenbach also identified problems with standard treatments of quantum mechanics, amongst them “causal anomalies”, and sought to remedy them by introducing a third truth value which he termed “indeterminacy”. However, there are important differences between the approaches. Firstly, Reichenbach’s work on three-valued logic was prior to Bell’s on non-locality, and to Reichenbach’s own later work on the principle of common cause. His discussion has nothing explicit to say about the problem of restoring the PCC in the face of Bell’s theorem, therefore. Moreover, the scheme given above is not a simple extension of Reichenbach’s ideas to solve the problem in hand; there are also significant technical

differences.

In the approach above, the normal Boolean algebra of events is retained, and only the normal logical operations are extended to the three-valued logic. In contrast, Reichenbach allows completely general logical operations, mapping the “input” truth values of events and pairs of events to the “outputs” in all possible combinations. For example, the logic employed above implies the rule $\tilde{\gamma}(A) = 1 \Rightarrow \tilde{\gamma}(A^c) = 0$, $\tilde{\gamma}(A) = 1/2 \Rightarrow \tilde{\gamma}(A^c) = 1/2$, $\tilde{\gamma}(A) = 0 \Rightarrow \tilde{\gamma}(A^c) = 1$. For Reichenbach this is only one of the possible logical operations with one argument. As well as this “diametrical negation” which he terms $-A$ he considers “cyclical negation” $\sim A$ that gives $\tilde{\gamma}(A) = 1 \Rightarrow \tilde{\gamma}(\sim A) = 1/2$, $\tilde{\gamma}(A) = 1/2 \Rightarrow \tilde{\gamma}(\sim A) = 0$, $\tilde{\gamma}(A) = 0 \Rightarrow \tilde{\gamma}(\sim A) = 1$ (our set theoretic notation, which implies a Boolean algebra of events, fails here, but hopefully the meaning remains clear). Likewise there are novel two-argument logical operations in Reichenbach’s approach. Every logical operation of this type is allowed. Thus, there are events in the three-valued logic corresponding to all the events at the meta-level described in section 5.1 (*i.e.* that have the same definite truth value as the given meta-level event). Given this, Reichenbach’s approach does not allow the strategy used above to give a causal explanation of the EPRB experiment: using the novel logical operations, an event can be defined in the past that is correlated with the setting events, violating our assumptions. Consider the model of the EPRB experiment given in section 4.5. Defining a logical operation “is definite”, $?A$, such that $\tilde{\gamma}(A) = 1/2 \Rightarrow \tilde{\gamma}(?A) = 0$, $\tilde{\gamma}(A) = 0 \text{ or } 1 \Rightarrow \tilde{\gamma}(?A) = 0$, we find that the event $?P_{A1}$ determines the setting A_s . Thus admitting Reichenbach’s extra logical operations leads to a version of the “meta-level” objection discussed in section 5.1. The answers to the objection are the same. Nothing forces us to allow these operations in our new logic; indeed, from the perspective given in the main text they appear to be an unnatural addition.

Another feature of Reichenbach’s approach is that it is explicitly operationalist in character. Thus he states “it is possible to introduce an intermediate truth value which may be called *indeterminacy*, and to coordinate this truth value to the group of statements which in the Bohr-Heisenberg interpretation are called *meaningless*” – that is, propositions about unmeasured properties. With this operational meaning for the third truth value, the principle of common cause that Reichenbach later formulated cannot be restored, as argued in section 3.2.

However, there are several points from Reichenbach’s discussion that do carry over. One is that complementarity can be given a meaning based on truth values in the logical system. Two events A and B may be considered complementary if and only if, for all $\tilde{\gamma} \in \tilde{\Theta}$, $\tilde{\gamma}(A) = 1/2$ if and only if $\tilde{\gamma}(B) \neq 1/2$. Thus events P_{A0} and P_{A1} in section 4.5 are complementary. This observation is based on a point made by Reichenbach, although in this context the definition of complementarity is expanded somewhat from its usual operational setting.

Reichenbach also recognises that introducing a third truth value may have an impact on the EPR “paradox”. After noting that Bohr’s opinions on EPR “do not seem to us to be stated sufficiently clearly to admit of an unambiguous interpretation”, Reichenbach confirms with EPR that “what is proved by the existence of correlated

systems is that it is not permissible to say that *the value of an entity before the measurement is different from the value resulting in the measurement* (added italics) – if, the context suggests, one assumes relativistic causal structure. Reichenbach’s proposed solution to the problem is that this italicised statement, though untrue, may not be false but instead hold the indeterminate truth value. However, it is not clear that Reichenbach’s proposal saves us from “causal anomalies”, as he calls them, because it is not made clear exactly what constitutes a causal anomaly in the EPR setting, especially once the three-valued logic is introduced. This clarification would have to wait for the later work of Reichenbach, and Bell, on causal principles. The three-valued definition of Einstein Locality given in section 4.3 provides such a definition for a three-valued logic. Using this definition, Reichenbach’s proposed solution is not adequate, as he does not provide causal antecedents for the inarguably determinate events: the measurement outcomes²⁰.

D More details of the EPRB experiment

This appendix gives some further explanation and discussion of points covered in section 4.5.

D.1 More on free settings

Before we come to discuss the details of the EPRB experiment, more needs to be said about the role of free settings. In the EPRB experiment there is more than one setting, and so questions arise beyond those already addressed in sections 2.2 and 4.3. For example, should logical combinations of settings also be treated as settings? And should events that are logical combinations of settings and other events require causal antecedents? We need to be more systematic about such questions to ensure that our model the EPRB experiment obeys freedom of settings and Einstein Locality for all events. We will start with the case of standard two-valued logic.

Setting events are a set of events in Σ . We will say that logical combinations of settings are also settings, *i.e.* the settings form a subalgebra of Σ (this may seem odd in that the “trivial events” Ω and \emptyset now qualify as settings, but this is the easiest way to proceed).

The main condition on settings is that they are not influenced by any past events in the model. Consider a setting A associated to region \mathcal{A} . If $A \cap \Theta$ is a proper subset of Θ (*i.e.* if there are histories in Θ for which A is false), then there can be no event C associated to $J^-(\mathcal{A}) \setminus \mathcal{A}$ such that $C \neq \emptyset$ and $C \cap \Theta \subset A \cap \Theta$. In other words C cannot be weakly correlated to A except in trivial cases. As well as this, settings associated to disjoint regions should not be weakly correlated to each other, a condition easily satisfied in the model of the EPRB experiment.

²⁰Perhaps the “meta-level version” of EL mentioned in section 5.1 would be more appropriate for Reichenbach’s logic, but again this runs into trouble if the truth “value of an entity” in a distant region can be changed from 1/2 to 1 or 0 by the free choice of an experiment.

As well as these conditions on settings, the causal condition of Einstein Locality is weakened for settings. This is because settings are supposed to be a “free choice”, and thus whatever it is that determines them need not be relevant for the rest of the model (for example, if we decided the settings based on what we had for breakfast, it should not be necessary to include our breakfast in the model to get a good description of the system in question). In order to exempt signals from Einstein Locality, we will weaken the second formulation of Einstein Locality given above.

As before we have a region \mathcal{X} , its past $\mathcal{R} = J^-(\mathcal{X})$ and its “exclusive past” $\mathcal{P} = J^-(\mathcal{X}) \setminus \mathcal{X}$ as shown in figure 1. The set $\Sigma|_{\mathcal{P}}$ is the subalgebra of events associated to the past region. Consider now the subalgebra Σ_{settings} generated by $\Sigma|_{\mathcal{P}}$ and the setting events associated to \mathcal{R} .

Einstein Locality (with settings): Consider a nonprobabilistic theory defined by $\{\Omega, \Sigma, \Theta, \mathcal{M}, \Delta\}$, and regions \mathcal{X} , $\mathcal{R} = J^-(\mathcal{X})$ and $\mathcal{P} = J^-(\mathcal{X}) \setminus \mathcal{X}$. The theory is *Einstein local* if the following condition holds for all regions \mathcal{X} :

$$\forall A \in \Sigma|_{\mathcal{R}}, \exists B \in \Sigma_{\text{settings}} \text{ such that } \gamma(A) = \gamma(B) \forall \gamma \in \Theta.$$

Previously, all causal antecedents had to be in the past region \mathcal{P} . Now we have expanded the allowed causal antecedents to also include setting events in the region \mathcal{R} itself. Thus settings are exempted in the sense that they are allowed to be their own causal antecedents, which is a trivial requirement. This also deals with events that are logical combinations of settings with other events. Consider for example the event $A \cap B$ where A is a free setting associated to region \mathcal{A} and B is a non-setting event associated to a spacelike region \mathcal{B} . Here, B itself requires a causal antecedent which we can call D , but $A \cap D$ is an acceptable causal antecedent for $A \cap B$, even though this event is not associated to $J^-(\mathcal{A} \cup \mathcal{B}) \setminus (\mathcal{A} \cup \mathcal{B})$.

As in the main text, these rules are easily extended to the case of theories using the three truth values. In this case, we still require the settings to form a subalgebra of Σ . To prevent past influence on settings, we require the following (which follows for the derived three-valued theory from the condition above for two-valued theories, in the same way as in section 4.3). Consider a setting A associated to region \mathcal{A} . If there exists $\tilde{\gamma} \in \tilde{\Theta}$ such that $\tilde{\gamma}(A) = 0$, then there can be no event C associated to $J^-(\mathcal{A}) \setminus \mathcal{A}$ such that, for all $\tilde{\gamma} \in \tilde{\Theta}$, $\tilde{\gamma}(C) = 0$ whenever $\tilde{\gamma}(A) = 0$, and that is nontrivial in the sense that there exists $\tilde{\gamma} \in \tilde{\Theta}$ such that $\tilde{\gamma}(C) = 1$. This is the three-valued version of the statement that C cannot be weakly correlated to A except in trivial cases.

Again, the statement of Einstein locality with settings is the same as the statement without settings, found in section 4.3 except with Σ_{settings} replacing $\Sigma|_{\mathcal{P}}$. These conditions – freedom of settings and Einstein locality with the exemption for settings – are the ones that the EPRB experiment must obey.

D.2 Application to the EPRB experiment

In the model of the EPRB experiment provided in the text, 16 basic three-valued truth valuations are provided for the events of the model: the setting events A_s and B_s , the results events A_r and B_r , and the four P variables assigned to the past. The following is a table showing the truth values for these events in each of these 16 truth valuations:

i	$\tilde{\gamma}_i(A_s)$	$\tilde{\gamma}_i(B_s)$	$\tilde{\gamma}_i(A_r)$	$\tilde{\gamma}_i(B_r)$	$\tilde{\gamma}_i(P_{A0})$	$\tilde{\gamma}_i(P_{A1})$	$\tilde{\gamma}_i(P_{B0})$	$\tilde{\gamma}_i(P_{B1})$
1	0	0	0	0	0	1/2	0	1/2
2	0	0	0	1	0	1/2	1	1/2
3	0	0	1	0	1	1/2	0	1/2
4	0	0	1	1	1	1/2	1	1/2
5	0	1	0	0	0	1/2	1/2	0
6	0	1	0	1	0	1/2	1/2	1
7	0	1	1	0	1	1/2	1/2	0
8	0	1	1	1	1	1/2	1/2	1
9	1	0	0	0	1/2	0	0	1/2
10	1	0	0	1	1/2	0	1	1/2
11	1	0	1	0	1/2	1	0	1/2
12	1	0	1	1	1/2	1	1	1/2
13	1	1	0	0	1/2	0	1/2	0
14	1	1	0	1	1/2	0	1/2	1
15	1	1	1	0	1/2	1	1/2	0
16	1	1	1	1	1/2	1	1/2	1

This table, however, does not fully specify the sixteen truth valuations $\{\tilde{\gamma}_i(\cdot)\}$. What, for example, is the value of $\tilde{\gamma}_1(P_{A1} \cup P_{B1})$? The answer to questions like these does not follow from the values in the truth table above as it would for classical logic. As explained in section 4.2, to specify these it is sufficient to specify a set of definite truth valuations from which the history is derived. Consider for example the following four definite histories:

i	$\gamma_i(A_s)$	$\gamma_i(B_s)$	$\gamma_i(A_r)$	$\gamma_i(B_r)$	$\gamma_i(P_{A0})$	$\gamma_i(P_{A1})$	$\gamma_i(P_{B0})$	$\gamma_i(P_{B1})$
1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	1
3	0	0	0	0	0	1	0	0
4	0	0	0	0	0	1	0	1

Indefinite truth valuations derived from subsets of this set can produce the truth values given for $\tilde{\gamma}_1(\cdot)$ above. To give $\tilde{\gamma}_1(P_{A1}) = \tilde{\gamma}_1(P_{B1}) = 1/2$ the subset must contain at least one definite truth valuation with each value for P_{A1} and P_{B1} . The sets $X = \{\gamma_1(\cdot), \gamma_4(\cdot)\}$ and $Y = \{\gamma_2(\cdot), \gamma_3(\cdot)\}$ have this property. These correspond to saying that, while the truth values given to P_{A1} and P_{B1} are indefinite, there are still definite statements to be made about their logical combinations. For example $\tilde{\gamma}_Y(P_{B1} \cup P_{A1}) = 1$ as this event has truth value 1 for both $\gamma_2(\cdot)$ and $\gamma_3(\cdot)$. The set

$Z = \{\gamma_1(\cdot), \gamma_2(\cdot), \gamma_3(\cdot), \gamma_4(\cdot)\}$ and all the three-member subsets also qualify, giving less definite events. As an example $\tilde{\gamma}_1(\cdot)$ can be set equal to $\tilde{\gamma}_Z(\cdot)$, and similarly, each other allowed history $\tilde{\gamma}_i(\cdot)$ can be derived from the set of four definite histories that share the definite truth values listed for $\tilde{\gamma}_i(\cdot)$ in the first table above.

However, we can see that several sets of truth valuations following the table above will satisfy our conditions of Einstein locality and freedom of settings. For the proof of freedom of settings, we will need to make the specification that for all sixteen of the indefinite truth valuations listed in the table, the conjunction of the two indefinite past events gives the same value w , which can be 0 or 1/2. In other words, $\tilde{\gamma}_k(P_{A1} \cap P_{B1}) = \tilde{\gamma}_l(P_{A1} \cap P_{B0}) = \tilde{\gamma}_m(P_{A0} \cap P_{B1}) = \tilde{\gamma}_n(P_{A0} \cap P_{B0}) = w$ for all values of the indices such that $1 \leq k \leq 4$, $5 \leq l \leq 8$, $9 \leq m \leq 12$ and $13 \leq n \leq 16$. Similarly, unique values x , y , and z must be given to the conjunctions of the indefinite past events and their compliments. That is, $\tilde{\gamma}_k(P_{A1} \cap P_{B1}^c) = \tilde{\gamma}_l(P_{A1} \cap P_{B0}^c) = \tilde{\gamma}_m(P_{A0} \cap P_{B1}^c) = \tilde{\gamma}_n(P_{A0} \cap P_{B0}^c) = x$ under the same conditions on the indices as before, and similarly for $P_{A1}^c \cap P_{B1}$ and $P_{A1}^c \cap P_{B1}^c$ and y and z . Specifying these four variables specifies all of the truth values for each truth valuation. While there are other specifications of the truth valuations consistent with the above table that would also obey our physical conditions, the main point of interest here is that even one exists.

The settings here are A_r and B_r but also all logical combinations of them such as $A_r \cap B_r$, $A_r \cap B_r^c$, and so on. Firstly, it will be shown that Einstein locality in the “settings version” holds. We do not require causal antecedents for the P events (or any non-trivial logical combinations with them) as they are the earliest events that we have modelled²¹). Thus we only need establish causal antecedents for events associated with the remaining region $\mathcal{A} \cup \mathcal{B}$. As mentioned in the main text, it can be seen from the table above that A_r and B_r have causal antecedents, because $\tilde{\gamma}((A_s \cap P_{A1}) \cup (A_s^c \cap P_{A0})) = \tilde{\gamma}(A_r)$ for all of the 16 basic histories, and similarly for B_r . The “causal antecedents” of the settings are themselves. All other events associated with $\mathcal{A} \cup \mathcal{B}$ are logical combinations of these events. The causal antecedent of any logical combination of such events simply the same logical combination of the causal antecedents of these events. This covers all events in the model, establishing Einstein Locality.

It remains to show freedom of settings. For each of the setting events, the potential weakly correlated “influencing” events are the events that lie to the past: the four P events and their logical combinations. There must be no event C associated to the past that has $\tilde{\gamma}_i(C) = 1$ for at least one value of i , and that also has $\tilde{\gamma}_i(C) = 0$ for all values of i such that $\tilde{\gamma}_i(A) = 0$, for any setting event A (apart from the trivial “setting” \emptyset). We only need show that these conditions do not hold for any “full specification” of the settings, for example $A_s \cap B_s$, $A_s \cap B_s^c$ and so on. If they do not hold for these events then they cannot be true of any disjunction of them, as a truth valuation cannot have $\tilde{\gamma}_i(A_1 \cup A_2) = 0$ if it is the case that $\tilde{\gamma}_i(A_1) = 0$ and $\tilde{\gamma}_i(A_2) = 0$.

²¹We can imagine a causal chain descending further into the past, but there is no need to model this. If there were correlations between some of these “initial events” that were associated to spacelike regions, leaving out their causal antecedents would be more questionable, but we do not have to address this issue here since all past events are associated to the same region.

Disjunctions of the full specifications cover all non-trivial logical combinations of the settings. Let us consider the full specification $A_s \cap B_s$, which is true for the truth valuations for which $13 \leq i \leq 16$ in the table and false for $1 \leq i \leq 12$.

Now let us consider full specifications of the past events, meaning the sixteen “finest-grained” events such that knowing that one of these events had truth value 1 would imply the truth value of all other past events. These are the events $P_{A0} \cap P_{A1} \cap P_{B0} \cap P_{B1}$, $P_{A0} \cap P_{A1} \cap P_{B0} \cap P_{B1}^c$ and so on. Let us label these events $\{F_j\}$ where j runs between one and sixteen (j is used to distinguish this index from the index i labelling the sixteen truth valuations, which is quite distinct). We will now show that, for any value of j , $\tilde{\gamma}_i(F_j) = 0$ when $1 \leq i \leq 12$ implies that $\tilde{\gamma}_i(F_j) = 0$ when $13 \leq i \leq 16$. Thus, it cannot be true that $\tilde{\gamma}_i(F_j) = 1$ for at least one value of i and that $\tilde{\gamma}_i(F_j) = 0$ for all values of i such that $\tilde{\gamma}_i(A_s \cap B_s) = 0$. In other words, F_j is not (non-trivially) weakly correlated to $A_s \cap B_s$.

The events $\{F_j\}$ involve the conjunction of two indefinite events. The truth value $\tilde{\gamma}_i(P_{A0} \cap P_{A1} \cap P_{B0} \cap P_{B1})$ is the same for $i = \{4, 8, 12, 16\}$: because of our assumptions on the truth value of conjunctions of indefinite events made above, it takes the value w . The truth value $\tilde{\gamma}_i(P_{A0} \cap P_{A1} \cap P_{B0} \cap P_{B1})$ also holds the same value, zero, for all other values of i : as can be seen from the table, there are zeros in the columns for some P events in all other values of i , which makes the conjunction definitely zero. A similar argument can be made for all fifteen other events in the set $\{F_j\}$. The result is that for all values of j , and for $1 \leq i \leq 4$, $\tilde{\gamma}_i(F_j) = \tilde{\gamma}_{i+4}(F_j) = \tilde{\gamma}_{i+8}(F_j) = \tilde{\gamma}_{i+12}(F_j)$. Moreover, it follows that this is true also for disjunctions of events in $\{F_j\}$, that is, for all non-trivial past events.

From this, it follows that, for all j , either $\tilde{\gamma}_i(F_j) = 0$ for all values of i , or it is non-zero for some i in the range $1 \leq i \leq 12$. Again, this is also true for disjunctions of events in $\{F_j\}$. This shows that none of the past events are (non-trivially) weakly correlated to the setting event $A_s \cap B_s$. The same argument can be adapted to the remaining three “full specification” setting events $A_s \cap B_s^c$, $A_s^c \cap B_s$ and $A_s^c \cap B_s^c$. From this, it follows that the freedom of settings condition holds.

References

- [1] J. Bell, *Physics* **1**, 195 (1964).
- [2] J. Bell, *Speakable and unspeakable in quantum mechanics* (CUP, Cambridge, 1987).
- [3] J. Bell, La nouvelle cuisine, in *Between science and technology*, pp. 216–233, Elsevier, 1990.
- [4] R. W. Spekkens, *Phys. Rev.* **A75**, 032110 (2007), quant-ph/0401052.
- [5] J. Butterfield, *British Journal for the Philosophy of Science* **43**, 41 (1992), pss/687884.

- [6] C. G. Timpson and H. R. Brown, Entanglement and relativity, in *Understanding Physical Knowledge*, edited by R. Lupacchini and V. Fano, Department of Philosophy, University of Bologna, 2002, quant-ph/0212140.
- [7] H. Reichenbach, (Berkeley: University of California Press, 1956), reissued 1991.
- [8] D. Bohm, Phys. Rev. **85**, 166 (1952).
- [9] D. Bohm, Phys. Rev. **85**, 180 (1952).
- [10] E. Nelson, Phys. Rev. **150**, 1079 (1966).
- [11] L. Smolin, (2006), quant-ph/0609109.
- [12] G. C. Ghirardi, A. Rimini, and T. Weber, Phys. Rev. D **34**, 470 (1986).
- [13] J. Bell, *Speakable and unspeakable in quantum mechanics* (CUP, Cambridge, 1987), chap. 22.
- [14] A. Kent, (2007), 0708.3710.
- [15] J. B. Hartle, Space-time quantum mechanics and the quantum mechanics of space-time, in *Proceedings of the Les Houches Summer School on Gravitation and Quantizations, Les Houches, France, 6 Jul - 1 Aug 1992*, edited by J. Zinn-Justin and B. Julia, North-Holland, 1995, gr-qc/9304006.
- [16] J. G. Cramer, Reviews of Modern Physics **58**, 647 (196).
- [17] P. Morgan, Journal of Physics A Mathematical General **39**, 7441 (2006), arXiv:cond-mat/0403692.
- [18] D. P. Rideout and R. D. Sorkin, Phys. Rev. **D61**, 024002 (2000), gr-qc/9904062.
- [19] G. Birkhoff and J. von Neumann, Annals of Mathematics **37**, 823843 (1936).
- [20] M. L. Dalla Chiara and R. Giuntini, *Quantum logics* (Kluwer, Dordrecht, 2002), p. 129228.
- [21] H. Reichenbach, (University of California Press, 1944), Reprinted by Dover 1998.
- [22] C. J. Isham and J. Butterfield, (1998), quant-ph/9803055.
- [23] Döring, A. and Isham, C. J., (2007), quant-ph/0703060.
- [24] R. D. Sorkin, J. Phys. **A40**, 3207 (2007), quant-ph/0610204.
- [25] R. D. Sorkin, (2010), 1004.1226.
- [26] F. Dowker and Y. Ghazi-Tabatabai, J. Phys. **A41**, 105301 (2008), 0711.0894.

- [27] J. Butterfield, *The British Journal for the Philosophy of Science* **58**, 805 (2007), arXiv:0708.2192.
- [28] J. Uffink, *Philosophy of Science* **66**, 525 (1999).
- [29] J. Henson, *Stud. Hist. Philos. Mod. Phys.* **36**, 519 (2005), quant-ph/0410051.
- [30] F. Selleri, *Foundations of Physics* **8**, 103 (1978), 10.1007/BF00708489.
- [31] J. Norton, *Philosophers Imprint* **3** (2003), to be reprinted in H. Price and R. Corry, *Causation and the Constitution of Reality*, Oxford: University Press.
- [32] N. D. Mermin, *Reviews of Modern Physics* **65**, 803 (1993).
- [33] T. Norsen, *Foundations of Physics Letters* **19**, 633 (2006), arXiv:quant-ph/0601205.
- [34] T. Norsen, *Foundations of Physics* **37**, 311 (2007), arXiv:quant-ph/0607057.
- [35] T. Placek, *British Journal for the Philosophy of Science* **51**, 445 (2000).
- [36] H. Putnam, Is logic empirical?, in *Boston Studies in the Philosophy of Science*, vol. 5, edited by R. Cohen and M. Wartofsky, pp. 216–241, Dordrecht: Reidel, 1968, Reprinted as The logic of quantum mechanics in H. Putnam, *Mathematics, Matter, and Method. Philosophical Papers*, vol. 1 (Cambridge: Cambridge University Press, 1975), pp. 174–197.
- [37] M. Dummett, Is logic empirical?, in *Contemporary British Philosophy*, 4th series, pp. 45–68, London: Allen and Unwin, 1976, Reprinted in M. Dummett, *Truth and other Enigmas* (London: Duckworth, 1978), pp. 269–289.
- [38] J. Hintikka, (Cornell University, 1962), Reprinted 2005.
- [39] N. Rescher and R. Brandom, *The logic of inconsistency* (Oxford, 1980).
- [40] M. Dummett, *Realism* (Harvard University Press, 1978), p. 145–165.
- [41] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Going Beyond Bell’s Theorem, in *Bell’s Theorem, Quantum Theory, and Conceptions of the Universe*, pp. 69–72, Kluwer, Dordrecht, 1989, 0712.0921.
- [42] N. D. Mermin, *American Journal of Physics* **58**, 731 (1990).
- [43] R. Coecke, B. Edwards, and R. W. Spekkens, (2010), arXiv:1003.5005.
- [44] J. Butterfield, *Int. Stud. Phil. Sci.* **21**, 295 (2007), 0708.2189.
- [45] A. Fine, *Do Correlations Need to be Explained?* (University of Notre Dame Press, 1989), pp. 175–94.
- [46] J. Bernstein, *Quantum profiles* (Princeton, 1991).